Ghent University
Faculty of Bioscience Engineering

Towards high spatial resolution air quality mapping: a methodology to assess street-level exposure based on mobile monitoring

ir. Joris Van den Bossche

Thesis submitted in fulfillment of the requirements for the degree of
Doctor (Ph.D) of Applied Biological Sciences

Academic year 2015-2016

Supervisors: Prof. dr. Bernard De Baets
Department of Mathematical Modelling, Statistics and Bioinformatics
Ghent University, Belgium

Prof. dr. ir. Dick Botteldooren
Department of Information Technology
Ghent University, Belgium

ir. Jan Theunis
Flemish Institute of Technological Research
VITO, Belgium

Examination committee:

Prof. dr. ir. Wolter Prins (Chairman)
Prof. dr. ir. Herman Van Langenhove
Prof. dr. ir. Timothy Van Renterghem
Prof. dr. ir. Sidharta Gautama
Prof. dr. Luc Int Panis
dr. Christoph Hüglin

Dean: Prof. dr. Marc Van Meirvenne

Rector: Prof. dr. Anne De Paepe

ir. Joris Van den Bossche

Towards high spatial resolution air quality mapping: a methodology to assess street-level exposure based on mobile monitoring

Thesis submitted in fulfillment of the requirements for the degree of

Doctor (Ph.D) of Applied Biological Sciences

Academic year 2015-2016

Dutch translation of the title:
Luchtkwaliteitskaarten met een hoge ruimtelijke resolutie: een methodologie om de blootstelling op straatniveau te beoordelen op basis van mobiele metingen

Please refer to this work as follows:
Joris Van den Bossche (**2016**). *Towards high spatial resolution air quality mapping: a methodology to assess street-level exposure based on mobile monitoring*, PhD Thesis, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Ghent, Belgium.

ISBN 978-90-5989-919-3

Dankwoord

Onderzoek is een teamsport, of zou het toch moeten zijn. En ik ben heel blij dat dit bij mij ook het geval was. Daarom wil ik in de eerste plaats mijn promotoren en collega's bedanken zonder wie dit doctoraat nooit was geweest wat het nu is geworden. Bernard, Dick, Jan, heel erg bedankt voor het mogelijk maken van dit doctoraat, voor de begeleiding en de altijd nauwgezette feedback. Ik heb hier veel van geleerd. Jan, als VITO-begeleider heb jij me het meest van dichtbij opgevolgd. En heb jij ook het meest met me afgezien. Jan, bedankt voor de vele interessante, diepgaande discussies die we hebben gevoerd en die me telkens tot meer inzicht brachten, en voor je geduld. Door mijn collega's van LKM op VITO (Jan P, Bart, Martine, in het eerste jaar ook Matteo, mijn bureaugenoot Guido, Jo en Rob voor de praktische zaken) had ik ook collega's had die echt met hetzelfde thema bezig waren, waar we samen aan projecten konden werken, waar met mijn resultaten iets gedaan werd. Dit was een heel grote meerwaarde voor mezelf en mijn doctoraat. En uiteraard wil ik ook VITO bedanken voor het financieel mogelijk maken van dit doctoraat. Verder wil ik nog een derde Jan bedanken voor de hulp bij de regressiemodellen.

Ik heb altijd zeer graag gewerkt zowel in Gent als op VITO. Daarvoor wil ik mijn collega's heel erg bedanken. De lunch- en koffiepauzes, de lekkere dessertjes, de voetbalmatchen, ... in Gent ga ik missen. Too many colleagues at UGent to sum up, but I want to specifically thank all the people who have worked in the simulation lab and my 'island' companions throughout the years. En specifiek ook open source en Python buddies Stijn en Timothy. Stijn mocht dan wel schrijven dat ik zijn grote voorbeeld was op dat vlak, maar je moet weten: het is eigenlijk andersom. Als er iemand is die het altijd in zijn eigen werk in de praktijk heeft proberen te brengen, is het Stijn wel.

Verder wil ik graag mijn ouders en broer en zus bedanken voor alle steun en me op alle mogelijke momenten binnen te laten stuiken in de Spillemansstraat. Mijn

huisgenoten van de lentestraat om daar altijd thuis te kunnen komen en terecht te kunnen. Mijn vrienden voor alle ontspannende momenten. De JNM voor de fantastische tijd en om me net op tijd oude sok te laten worden zodat ik mijn doctoraat kon afwerken.

Ik ben ook zeer dankbaar voor de mogelijkheden die dit doctoraat me hebben geboden, om mezelf te kunnen ontplooien en verschillende wegen te verkennen. De twee jobs die ik op dit moment doe, zowel rond luchtkwaliteit als in open source, waren beide niet mogelijk geweest zonder dit doctoraat. Daar ben ik zeer dankbaar voor, en benieuwd naar wat de toekomst zal brengen.

Joris

Summary

Background and objectives

There is extensive evidence that exposure to air pollution has both acute and chronic effects on human health. However, pollutant concentrations within the urban environment are highly variable in space and time, and thus the personal exposure of people can also differ considerably. Especially traffic-related pollutants, such as NOx, black carbon (BC) and ultrafine particles (UFP), show a high variability in concentrations at the street level. Traditionally, exposure to air pollution is evaluated with measurements at fixed monitoring stations. However, these measured concentrations are not always representative for the surrounding area. The use of portable devices makes it possible to perform mobile measurements and to generate additional data at locations where stationary measurements are lacking. As such, mobile monitoring can be used to map the variability in street-level concentrations people are actually exposed to.

Mobile monitoring is increasingly used and can lead to air quality data at a high spatial resolution, yet, at the expense of the temporal coverage. Given the high temporal variability of urban air quality and the limitations in temporal coverage of mobile monitoring, the methodology can be questioned and evaluations thereof are lacking in literature. In this thesis, the potential of mobile monitoring to map street-level exposure to urban air pollution at a high spatial resolution is assessed. Tools and methods to gather, process and extrapolate such data will be developed and evaluated.

Material and methods

To evaluate the potential of mobile monitoring methodologies, two measurement campaigns were set up in the city of Antwerp, Belgium. The first campaign is a targeted mobile monitoring campaign. In this campaign, two predefined routes (2

and 5 km long) were repeatedly monitored, and a total of 256 and 96 runs along the two routes were obtained over a period of 4 weeks. The second campaign is an opportunistic mobile monitoring campaign. City guards measured BC during their daily routines over a period of one year. As a consequence of their working routines, there is no control over the followed route and the campaign has an unstructured set-up.

The opportunistic measurement campaign is further used to develop land use regression models to obtain pollution maps with a more complete spatial and temporal coverage. Different regression techniques, both linear and non-linear, are evaluated for a spatial model, and finally a spatio-temporal model is built to estimate trip exposure.

Results and conclusions

Based on the extensive targeted campaign, it is shown that the mobile measurements converge to a typical pattern when using a carefully developed methodology. This shows that mobile monitoring is a suitable approach for mapping the urban air quality at a high spatial resolution, and can provide insight into the spatial variability at a resolution which would not be possible with stationary monitors. However, a large number of repeated measurements are still required to obtain representative results. A careful set-up is needed with a sufficient number of repetitions in relation to the desired reliability and spatial resolution, depending on the aim of the campaign. Specific data processing methods such as background normalisation and event detection are applied.

A possible way to gather a large number of measurements is to make use of existing mobile infrastructure or people's common daily routines to move measurement devices around, which we defined as opportunistic measurements. With the collaboration of the city wardens of Antwerp, we were able to gather a unique dataset showing the potential of opportunistic mobile monitoring. The opportunistic approach can identify broad spatial trends of the urban air quality over a wider area, enabling applications including hotspot identification, personal exposure studies, regression mapping, etc. But, the results also emphasize the need for repeated measurements and careful processing and interpretation of the data. This study also shows the potential of participatory and crowdsourcing campaigns to gather air quality data.

Mobile monitoring typically does not result in city-wide pollution maps with a full spatial and temporal coverage. Based on the data from the opportunistic measurement campaign, LUR models are built to predict the street-level exposure to BC.

Different types of models (both linear and non-linear) are compared and evaluated in an as opposed to many studies in literature accurate manner using spatial cross-validation. For the first time, a spatio-temporal model was developed that can predict dynamic trip exposure based on opportunistic measurements. Exposure estimates were more accurate than estimates based on a fixed site monitoring station and a spatial LUR model, and can be used to estimate exposure of cyclists or pedestrians to traffic-related pollution based on only a GPS track. The results demonstrated the potential of the spatio-temporal LUR model as a first step towards a real-time dynamic pollution map that can be updated based on real-time measurements. Such real-time pollution maps can be used for providing personalized information about air quality to citizens.

Samenvatting

Achtergrond en doelstellingen

Blootstelling aan luchtverontreiniging heeft een duidelijke, negatieve impact op de menselijke gezondheid. In de stedelijke omgeving zijn de concentraties van verschillende polluenten echter zeer variabel in ruimte en tijd, en de individuele blootstelling kan dus ook aanzienlijk verschillen. Specifiek verkeersgerelateerde polluenten, zoals NOx, zwarte koolstof (black carbon, BC) en ultrafijne deeltjes (ultrafine particles, UFP), vertonen een grote variabiliteit in concentraties op straatniveau. Traditioneel wordt de blootstelling aan luchtverontreiniging geëvalueerd op basis van metingen in stationaire meetstations van de overheid. Deze meetstations geven zeer accurate metingen, maar door hun beperkte verspreiding zijn ze niet altijd representatief voor de hele stedelijke omgeving. Het gebruik van draagbare meettoestellen maakt het mogelijk om mobiele metingen uit te voeren en data te verzamelen op plaatsen waar stationaire metingen ontbreken. Op deze manier kan mobiel meten gebruikt worden om de variabiliteit in de concentraties op straatniveau in kaart te brengen.

Mobiel meten wordt steeds meer gebruikt en kan resulteren in luchtkwaliteitsdata met een hoge ruimtelijke resolutie. Maar, deze ruimtelijke resolutie wordt bereikt ten koste van de dekking van de metingen in de tijd. Gelet op de sterke temporele variabiliteit in de stedelijke luchtkwaliteit en deze beperkte dekking in de tijd, kan de methode van mobiel meten in vraag worden gesteld. Bovendien ontbreken evaluaties hiervan in de wetenschappelijke literatuur. In deze thesis wordt het potentieel beoordeeld van mobiel meten met als doel de blootstelling aan luchtverontreiniging in kaart te brengen met een hoge ruimtelijke resolutie. Methoden om dergelijke data te verzamelen, verwerken en extrapoleren werden ontwikkeld en geëvalueerd.

Materiaal en methoden

Om het potentieel van mobiel meten te evalueren werden twee meetcampagnes opgezet in Antwerpen, België. De eerste campagne was een specifiek geplande, mobiele meetcampagne. Twee vooraf bepaalde routes (van 2 en 5 km lang) werden herhaaldelijk gevolgd, en in totaal werden 256 en 96 herhalingen voor de twee routes uitgevoerd gespreid over een periode van 4 weken. De tweede campagne was een mobiele meetcampagne waarbij stadswachten in Antwerpen gedurende een jaar de BC concentraties hebben gemeten tijdens hun dagelijkse toezichtrondes. Als gevolg van hun werkroutines was er geen controle over de gevolgde route.

De tweede meetcampagne is verder gebruikt om landgebruiksregressiemodellen (land use regression, LUR) op te stellen. Deze regressiemodellen kunnen gebruikt worden om een luchtkwaliteitskaart met een meer volledige dekkingsgraad in de ruimte en tijd te bekomen. Verschillende regressietechnieken, zowel lineaire als niet-lineaire, werden geëvalueerd voor een spatiaal model, en uiteindelijk werd een spatio-temporeel model ontwikkeld om de blootstelling tijdens een verplaatsing in te schatten.

Resultaten en conclusies

De resultaten van de eerste, specifiek geplande meetcampagne tonen aan dat de mobiele metingen convergeren naar een typisch patroon wanneer de data zorgvuldig worden verwerkt. Mobiel meten kan bijgevolg een geschikte aanpak zijn om de stedelijke luchtkwaliteit in kaart te brengen met een hogere ruimtelijke resolutie dan mogelijk is met stationaire metingen. Een groot aantal herhaalde metingen blijft echter noodzakelijk om representatieve resultaten te bekomen. Een zorgvuldige opzet van de meetcampagne met voldoende herhalingen, afhankelijk van de gewenste betrouwbaarheid en ruimtelijke resolutie, is nodig. Specifieke dataverwerkingsmethoden zoals achtergrond normalisatie en detectie van occasionele pieken werden toegepast.

Een mogelijke manier om dit groot aantal herhaalde metingen te verzamelen is door gebruik te maken van bestaande infrastructuur of dagelijkse routines om de meettoestellen rond te bewegen in de stad. Dit definiëren we als opportunistische metingen. In samenwerking met de stadswachten van Antwerpen werd een unieke dataset verzameld die het potentieel van opportunistisch mobiel meten aantoont. De opportunistische aanpak kan duidelijke ruimtelijke trends in de luchtverontreiniging in kaart brengen voor een groter gebied, wat toepassingen zoals het identificeren van hotspots, persoonlijke blootstellingsstudies, regressiemodellen, enz.

mogelijk maakt. Maar, de resultaten benadrukken ook de nood aan herhaalde metingen en een zorgvuldige verwerking en interpretatie van de data. Deze studie toont verder het potentieel aan van participatieve methoden om luchtkwaliteitsdata te verzamelen.

Mobiel meten leidt typisch nog niet tot een kaart van de hele stad met een volledige ruimtelijke en temporele dekking. Op basis van de data van de opportunistische meetcampagne werden LUR modellen opgesteld om de blootstelling aan BC op straatniveau te voorspellen op plaatsen waar er niet werd gemeten. Verschillende model-types (zowel lineair als niet-lineair) werden vergeleken en geëvalueerd op een - in tegenstelling tot de gebruikte methodes in veel studies in de literatuur - accurate manier door gebruik te maken van ruimtelijk gestratificeerde cross-validatie. Een spatio-temporeel model werd ontwikkeld dat de blootstelling tijdens verplaatsingen kan voorspellen op basis van opportunistische metingen. De voorspellingen van dit model waren beter dan voorspellingen op basis van een stationair meetstation of op basis van een spatiaal LUR model, en kunnen gebruikt worden om de blootstelling van voetgangers en fietsers aan verkeersgerelateerde luchtverontreiniging in te schatten op basis van hun GPS track. De resultaten tonen het potentieel aan van het spatio-temporele LUR model als een eerste stap naar real-time dynamische luchtkwaliteitskaarten die telkens kunnen gepdatet worden op basis van real-time metingen. Dergelijke luchtkwaliteitskaarten kunnen gebruikt worden om gepersonaliseerde informatie over luchtkwaliteit te bezorgen aan burgers.

Contents

List of Abbreviations

AQG Air Quality Guideline

BC Black Carbon

CV Cross-Validation

EBC Equivalent Black Carbon

EC Elemental Carbon

EEA European Environment Agency

EPA US Environmental Protection Agency

EU European Union

EV Explained Variance

GAM Generalized Additive Model

GIS Geographic Information System

GPS Global Positioning System

HV Hold-out Validation

IARC International Agency for Research on Cancer

LASSO Least Absolute Shrinkage and Selection Operator

LOOCV Leave-One-Out Cross-Validation

LUR Land Use Regression

MAAP Multi-Angle Absorption Photometry

MAC Mass-specific Absorption Cross-section

OC Organic Carbon

OSM OpenStreetMap

PM Particulate Matter

RMSE Root Mean Square Error

SVF Sky View Factor

SVR Support Vector Regression

UFP Ultrafine Particles

VITO Flemish Institute for Technological Research

VMM Flemish Environmental Agency

VOC Volatile Organic Compounds

WHO World Health Organization

CHAPTER 1

Problem statement, research objectives and outline

"There is an old joke about a statistician who drowns in a river of on average half a meter deep. The same joke could apply to someone who wants to test the European NO_2 limits and uses a spatial average over a complete region, although it is questionable whether much laughter would be drawn from the public."

- Lefebvre et al. (2013)

1.1 Introduction

Exposure to air pollution has both acute and chronic effects on human health and can lead to respiratory and heart diseases, lung cancer and reduced life expectancy. Especially in urban areas, where most of the European population lives and which are typically hotspots of air pollution, a lot of people are exposed to air pollution. Important differences in pollutant concentrations occur over the day and between different micro-environments in a city. Especially traffic-related pollutants such as NOx, ultrafine particles (UFP) and black carbon (BC) show a high variability in concentrations at the street level. These highly variable concentrations result in considerable differences in personal exposure to air pollution. Traditionally, exposure to air pollution is evaluated with measurements at fixed monitoring stations. However, these measured concentrations are not always representative for the surrounding area. Therefore, fixed monitoring stations were found to be relatively poor predictors of exposure concentration levels experienced by individuals

in the urban transport micro-environment. The ability to measure air quality at a higher spatial and temporal resolution can yield a tremendous advance in understanding variability of pollutants in urban environments and their association to health effects, and thus the ability to take the most appropriate and effective measures.

Based on the current monitoring practice, the available data on air pollution are often insufficient to accurately estimate the actual exposure of citizens. A better assessment of the spatial and temporal variability is hampered by the cost of the measurements. But, advances in sensor technology and the emergence of portable and lower-cost sensing devices give rise to new opportunities for air quality monitoring. With the use of portable devices, mobile measurements can be performed. Mobile monitoring allows to acquire air quality data at a higher spatial resolution than traditional fixed monitoring stations. As such, mobile data can be used to assess the variability in concentration levels within the urban environment and to map the street-level concentrations people are actually exposed to at specific locations. When carried by individuals, mobile devices can be used to directly measure personal exposure. Smaller and lower-cost sensor devices can potentially be deployed in larger quantities and form networks of sensors, and in this way also increase the spatial coverage of the measurements. Furthermore, with portable devices it becomes possible to engage individual volunteers equipped with the sensors in a participatory or community-based monitoring strategy. This has the potential to gather the large amount of data needed to more accurately estimate the actual exposure of citizens.

1.2 Problem Statement

The ability to obtain air quality data with a high spatial resolution is needed for several reasons. In the first place, it will allow for a better understanding of the concentration variability in a complex urban environment and of human exposure to urban air pollution. It can be used to assess the spatial variability, identify hotspots and investigate the influence of local sources. Such information will lead to a better underpinning of policy measures. Combined with activity patterns and activity-based models, it can lead to a more accurate assessment of exposure to air pollution both at the individual and at the population level. Higher resolution air quality data will provide a better picture to assess whether the regulations on air pollution levels are met, and will enable to better inform and engage the public.

Several strategies can be pursued to obtain data and pollution maps with a high spatial resolution. These strategies include the deployment of a dense network of stationary low-cost sensors, mobile monitoring or modelling techniques. Models can be both empirical models based on collected data and dispersion models. However, dispersion models are complex, need very detailed input data and do often not result in street-level concentrations. More detailed dispersion models that include small-scale variability are not available for all cities. Currently, to be able to characterize the small-scale variability in all sorts of environments, measurements are needed.

However, the new opportunities in air quality sensing also give rise to new methodological challenges. Many of the currently available low-cost sensors experience a variety of problems regarding sensitivity, selectivity and stability. At this moment, these sensors cannot yet be used as such to measure outdoor air quality. Other portable monitors with better performance exist for components such as ultrafine particles and black carbon, but are still quite expensive. An example is the microaethalometer, a portable monitor for BC. Mobile monitoring can lead to air quality data at a high spatial resolution, yet, at the expense of the temporal coverage. Due to the high temporal variability of the urban air quality and the mobile nature of the measurements, the representativeness of the mobile measurements is a major issue. Methods and guidelines to collect and process these data in a proper way have to be drafted. How to integrate the new data in modelling frameworks has to be explored.

1.3 Objectives of this research

In this thesis, some of the above-mentioned challenges are investigated. More specifically, this thesis focuses on the challenges related to mobile monitoring to map the urban air quality at a high spatial resolution: the development and evaluation of methods to gather, process and extrapolate such data. This thesis focuses on the average street-level exposure to traffic-related pollution in the urban environment, i.e. to UFP and BC. In this context, average street-level exposure refers to the concentration levels people are typically exposed to at specific street locations, i.e. the average concentration at a specified location representative for a specified time period. It does not refer to the integral personal exposure, integrated over the micro-environments the individual has visited during a certain time period. It also does not refer to the inhaled dose, i.e. not taking into account the physical activity and inhalation rate. This thesis therefore focuses on charac-

terising one of the micro-environments required to estimate the integral personal exposure.

More specifically this thesis focuses on the following objectives:

1. **The evaluation of measurement devices for mobile applications.**

   In absence of reliable low-cost sensors, this thesis will focus on a higher quality device. More specifically, the micro-aethalometer, a portable black carbon monitor, will be evaluated for use in mobile monitoring based on a meta-analysis of existing campaigns.

2. **The development of methodologies for mobile air quality monitoring to map street-level exposure.**

   Different monitoring strategies for data collection, both targeted and opportunistic sampling, will be explored, their potential to map the average street-level exposure to urban air pollution will be critically assessed and data processing methods will be proposed.

3. **The extrapolation of measurements to draw an average street-level exposure map with a high spatial and/or temporal resolution.**

   More specifically, land use regression (LUR) models will be developed based on mobile measurements. The use of mobile measurements will be evaluated and different models will be compared using a custom cross-validation scheme. In a second step, the development of dynamic spatio-temporal exposure maps is explored, and their use for street-level exposure assessment is evaluated.

1.4 Related research activities at VITO and UGent

This thesis is strongly embedded in current research projects at VITO. It builds further upon the know-how and experience related to mobile monitoring (Berghmans et al., 2009; Van Poppel et al., 2013; Elen et al., 2013; Peters et al., 2013b) and the use of the micro-aethalometer and personal exposure monitoring (Dons et al., 2011, 2012, 2013a). In the PhD thesis of Evi Dons (Dons, 2013), the feasibility of personal monitoring was demonstrated. Personal monitoring revealed valuable insights into the exposure to BC during different activity types. In my PhD thesis, the mobile measurements are not used to look at personal exposure of the carriers, but to map the spatial variability at high resolution. The third objective of the extrapolation of the measurements builds further upon the work

of Evi Dons on LUR models, in which she built LUR models based on stationary measurements to include in activity-based exposure models. In my PhD thesis, mobile measurements with a higher spatial resolution are used instead of stationary measurements, a temporal component to obtain dynamic models is added, and non-linear regression techniques are evaluated.

The objectives of my PhD thesis are also strongly related to the research of Luc Dekoninck at UGent (Dekoninck, 2015). In his thesis, he uses noise as a proxy for traffic-related air pollution (BC) and builds spatio-temporal models for personal exposure to traffic-related particulate matter based on instantaneous mobile noise measurements (Dekoninck et al., 2013, 2015). Both theses share similar objectives and monitoring methodologies (mobile monitoring), but this thesis uses a different approach by focusing on directly measuring the air pollution instead of using noise as a proxy and by using opportunistic measurements. The spatio-temporal model developed in the underlying PhD thesis is applicable without instantaneous measurements and goes beyond the initial steps taken by Dekoninck in this direction. The advantages and disadvantages of both approaches will be discussed in Sections 10.4.3 and 11.4.3.

1.5 Outline of the thesis

This thesis consists of three main parts, following the three research objectives. The first part focuses on air quality and air quality monitoring tools. The first chapter (Chapter 2) describes the broader context of air pollution and personal exposure to air pollution, and then investigates new trends in air quality monitoring tools and methods. Chapter 3 focuses on one specific monitoring device to measure black carbon, the micro-aethalometer. This is a portable device that can be used for mobile monitoring, and in this chapter an extensive validation of this device is presented.

The second part of this thesis deals with mobile monitoring. The first chapter (Chapter 4) describes the context and provides a framework for mobile monitoring. A targeted, carefully designed mobile measurement campaign is presented in the next chapter (Chapter 5) and the high variability in UFP and BC concentrations is described. In Chapter 6 the mobile monitoring methodology is elaborated based on the same campaign. Different aspects of the data processing including spatial aggregation, dealing with peak concentrations and temporal variations and the requirement for repeated measurements are discussed. In contrast to the targeted monitoring campaign of the previous two chapters, mobile measurements

can also be gathered in an opportunistic setting, taking advantage of existing mobile infrastructure or people's common daily routines. Chapter 7 presents a case study based on such opportunistic measurements and explores the potential of this approach. Finally, Chapter 8 summarizes the practical considerations and guidelines for mobile monitoring campaigns, and briefly presents the results of some applications.

The third part of this thesis concerns the extrapolation of measurements. The first chapter (Chapter 9) gives an overview of the different methods to obtain pollution maps, with special focus on land use regression (LUR) modelling. The development of a spatial LUR model based on mobile measurements is investigated in Chapter 10. Different LUR model building techniques are explored and attention is paid to the implementation of a proper cross-validation to ensure a valid assessment of the predictive ability of the model. Chapter 11 builds further upon these results and presents the development of a spatio-temporal LUR model and its application for street-level exposure assessment.

The final chapter provides general conclusions and perspectives (Chapter 12).

1.6 Publications based on this thesis

Several chapters from this thesis have been published or have been submitted in a modified form:

- Chapter 5: Peters, J., Van den Bossche, J., Reggente, M., Van Poppel, M., De Baets, B. and Theunis, J. (2014). Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. *Atmospheric Environment*, 92 31-43.
- Chapter 6: Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J. and De Baets, B. (2015) Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. *Atmospheric Environment*, 105 148-161.
- Chapter 7: Van den Bossche, J., Theunis, J., Elen, B., Peters, J., Botteldooren, D., and De Baets, B. (2016). Opportunistic mobile air pollution monitoring: a case study with city wardens in Antwerp. *Atmospheric Environment*, 141 408-421.
- Chapter 10: Van den Bossche, J., Theunis, J., Verwaeren, J., Botteldooren, D., and De Baets, B. (2016). Development of land use regression models for black carbon based on mobile measurements in the urban environment. *Environmental Science & Technology* (submitted).

The section on the EveryAware project in Section 2.3.2 is based on the following publication:

- Sirbu, A., Becker, M., Caminiti, S., De Baets, B., Elen, B., Francis, L., Gravino, P., Hotho, A., Ingarra, S., Loreto, V., Molino, A., Mueller, J., Peters, J., Ricchiuti, F., Saracino, F., Servedio, V., Stumme, G., Theunis, J., Tria, F., Van den Bossche, J. (2015). Participatory patterns in an international air quality monitoring initiative. *PLOS ONE*, 10(8), e0136763.

Furthermore, other chapters (including Chapter 3 and 11) will be submitted for publication.

PART I

AIR QUALITY MONITORING

CHAPTER 2

New approaches in outdoor air quality monitoring

2.1 Outdoor air quality

Ambient air pollution is estimated to cause 3.7 million deaths each year worldwide (WHO, 2014). Also in Europe, air pollution is the single largest environmental health risk, responsible for more than 430,000 premature deaths (EEA, 2015a). Currently, around 75 % of the population in the European Union lives in urban areas, which are typically hotspots of air pollution (EEA, 2013a).

2.1.1 Air pollution

The air around us contains a complex mixture of gases and particles, affected by emissions from different sources, both anthropogenic and natural. Pollutants can be emitted directly into the atmosphere (primary pollutants), are dispersed by wind, and new pollutants can be formed by physical or chemical reactions in the atmosphere (secondary pollutants). Pollutants of major public health concern include particulate matter (PM), including ultrafine particles (UFP) and black carbon (BC), ground level ozone (O_3) and nitrogen dioxide (NO_2) (WHO, 2006; Hoek et al., 2010; Patel et al., 2010; Pope et al., 2009). But also other pollutants, such as metals and polycyclic aromatic hydrocarbons (PAHs), are linked to adverse health effects. In the indoor environment, some additional pollutants such as formaldehyde and radon are important as well (WHO, 2010), but this thesis will further focus on outdoor air pollution.

A prominent example of a secondary pollutant is ground level O_3, a major constituent of photochemical smog. It is formed by the reaction of pollutants such as nitrogen oxides (NO_x) and volatile organic compounds (VOC) with sunlight (photochemical reaction). Nitrogen oxide (NO_x) consists of a mixture of nitrogen dioxide (NO_2) and nitrogen monoxide (NO). Combustion processes at high temperatures initially emit NO, which has only a short lifetime in the atmosphere. NO is converted into NO_2 through reactions with oxygen and ozon. NO_2 has a longer lifetime and it is mainly NO_2 that is harmful for the environment. NO_x can be seen as a brown haze dome above cities (VMM, 2015).

Particulate matter (PM) is a complex mixture of small particles and liquid droplets suspended in the atmosphere, and vary in origin, size, shape and chemical composition. Primary particles are directly emitted from combustion or natural processes, whereas secondary particles are formed in the atmosphere from precursor compounds such as NH_3, NO_x, SO_2 and VOCs (Seinfeld and Pandis, 2006; Heal et al., 2012). Aerodynamic PM diameters are in the 0.01 - 100 μm range (Figure 2.1). The distribution as a function of particle size is typically characterised by three modes. The modes reflect the dominant processes giving rise to ambient PM. Particles smaller than $\sim$50 nm are termed the nucleation mode and are produced by condensation within the atmosphere or in the exhaust emissions from combustion processes. Nucleation mode particles are short-lived (minutes to hours) and grow by coagulation to form the accumulation mode ($\sim$50 nm to 1 μm). The coarse-mode particles ($>$ 1 μm) are mostly primary particles generated by mechanical processes (Heal et al., 2012). The total mass concentration is mainly determined by coarse particles and PM in the accumulation mode. Ultrafine particles ($<$ 0.1 μm aerodynamic diameter) have very small mass but contribute most to the particle number concentration.

Measurement of PM is mostly based on the mass concentration (in $\mu g\, m^{-3}$). Different classes based on its size are distinguished: the mass of all particles with an aerodynamic diameter below 10 μm (PM_{10}) or below 2.5 μm ($PM_{2.5}$). Those metrics are used in regulation. Additionally, also black carbon and ultrafine particles are distinguished as components of PM.

Ultrafine particles

Ultrafine particles (UFP) are generally defined as particles having a diameter of less than 0.1 μm. Typically, they are characterized by the particle number concentration (PNC), as their mass is negligible. Motor vehicle emissions constitute the major source of UFPs in urban environments (Morawska et al., 2008). UFPs

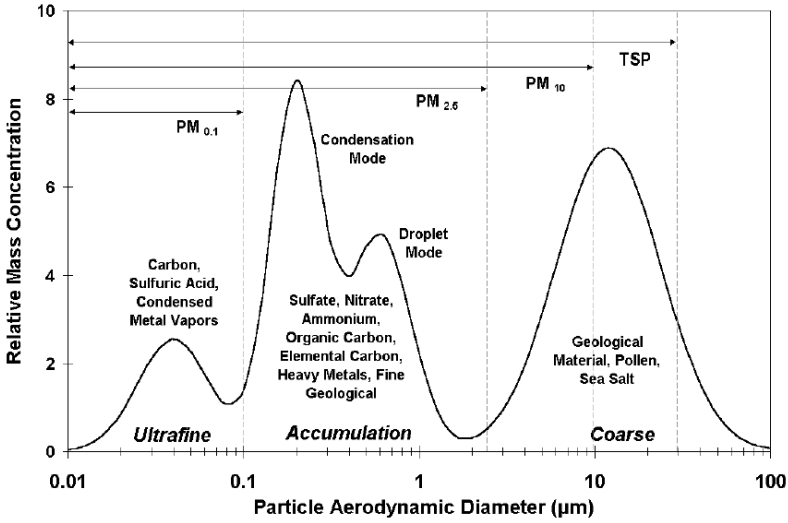


Figure 2.1: Relative mass concentration in function of the particle aerodynamic diameter.

have a transient nature with short lifetimes and rapidly grow through atmospheric processes of coagulation and/or condensation to larger complex aggregates (Seinfeld and Pandis, 2006; Kumar et al., 2010). Because of the short atmospheric lifetime of UFPs, the regional background is less important than for fine particles. Therefore, UFPs have a higher spatial variation than $PM_{2.5}$ (WHO, 2013; Reche et al., 2011; Klompmaker et al., 2015). This will be discussed in more detail in Section 2.1.4.

Black carbon

Black carbon (BC) is a constituent of fine particles. It is the carbon-containing component of PM that absorbs light at all wavelengths. PM also contains carbonaceous components that will absorb light less strongly, which is typically called organic carbon (OC). Black carbon is a product of incomplete combustion, and it is also referred to as soot. Major sources are traffic (particularly diesel engines), residential heating and open biomass burning, including forest fires. BC has a short atmospheric residence time of days to weeks (EPA, 2012; EEA, 2013b). BC is a good indicator of combustion-related air pollution, and exposure to BC is recognized by the World Health Organisation to be associated with cardiovascular mortality (Janssen et al., 2012).

BC is a useful qualitative description when referring to light-absorbing carbon-containing substances in atmospheric aerosol. However, for quantitative applications the term requires clarification on how the quantity is measured (Petzold et al., 2013). Petzold et al. (2013) proposes the term equivalent black carbon (EBC) to denote black carbon measurements derived from optical methods. Elemental carbon (EC) should be used when referring specifically to the carbon content of carbonaceous matter.

Sources

The outdoor air quality is affected by emissions from different sources, such as traffic, industry, agriculture and households. The contribution of the different sources depends on the pollutant and the region. The emissions of several primary pollutants in Flanders are shown in Figure 2.2. The transport sector is the largest contributor to NO_2 emissions. Both BC and NO_2 are mainly emitted by the combustion of fossil fuels in motor vehicles and stationary sources (residential heating, power generation). Especially in the urban environment, traffic is a major contributor to the local air pollution. Traffic also causes non-combustion emissions, due to resuspension of road dust, tire wear and brake wear (HEI, 2010).

Primary emissions of PM are mainly caused by households (residential heating) and industry. But, when including secondary emissions, also agriculture is a main source of PM (mainly due to the emission of NH_3 from livestock and fertilizers that lead to secondary PM formation). Also worldwide, agriculture is an important contributor to PM Lelieveld et al. (2015). Wood burning was found to contribute up to 10 % of PM_{10} in cities of Northwest Europe (VMM, 2013; Fuller et al., 2013).

Also natural sources contribute to air pollution. In Europe, the main natural sources of atmospheric aerosols are African dust (windblown desert dust), sea spray and wildfires. Volcanic eruptions can affect air quality as well (Viana et al., 2014).

2.1.2 Impact of air pollution on human health and the environment

Health effects of exposure to air pollution

Air pollution is the single largest environmental health risk in Europe and the world, and has a substantial disease burden (EEA, 2015a; Lim et al., 2012; WHO,

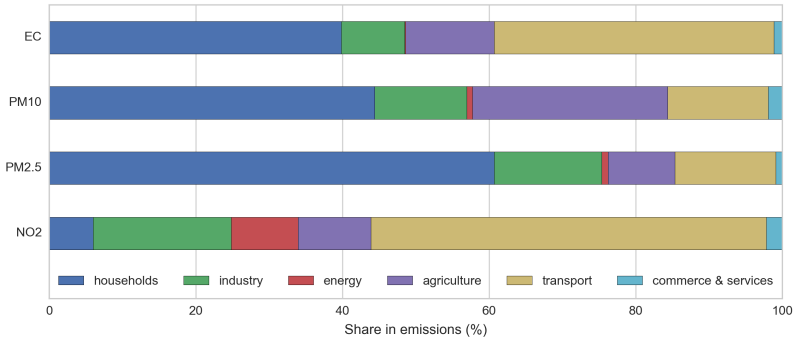


Figure 2.2: The share in emissions for different sectors in Flanders in 2013 (Data source: http://www.milieurapport.be).

2014). Air pollution is associated with many health impacts. The WHO estimates that about 80 % of outdoor air pollution-related premature deaths are due to heart disease and strokes. Air pollution is further associated with lung diseases such as chronic obstructive pulmonary disease (COPD) or acute lower respiratory infections (14 % of premature deaths) and lung cancer (6 % of premature deaths) (WHO, 2014; Lelieveld et al., 2015). The recent REVIHAAP (Review of evidence on health aspects of air pollution, WHO 2013) study concluded that the evidence has only strengthened over the last years. In addition to causing premature death, air pollution increases the incidence of a wide range of diseases (e.g. respiratory and cardiovascular diseases and cancer) and is associated with reduced foetal growth and pre-term birth in prenatally exposed children (WHO, 2013; EEA, 2015a). Outdoor air pollution has been classified as carcinogenic (IARC, 2015).

In terms of the potential to harm human health, PM is one of the most important air pollutants (Lim et al., 2012; Apte et al., 2015). The small particles can penetrate deep into the lungs and sensitive regions of the respiratory system, and cause a wide range of health problems (mainly cardiovascular and respiratory disorders). The diversity of particle size and chemical composition considerably complicates pinpointing the specific causal associations between exposure to particles, adverse human health effects and the contribution of different sources to ambient PM (Heal et al., 2012). Not all components of PM are equally harmful. In particular, long-term exposure to transition metals in ambient PM, originating from traffic and industry, may be more strongly related to the health effects of PM than the mass concentration (Hampel et al., 2015).

NO_2 and O_3 are also strongly related to health effects. Exposure to O_3 is mainly related to respiratory health problems, causing breathing problems, triggering

asthma, reducing lung function and causing lung diseases (WHO, 2013). NO_2 is a toxic gas that causes inflammation of the airways at short-term exposure to concentrations exceeding 200 $\mu g\,m^{-3}$. Epidemiological studies have shown that reduced lung function and symptoms of bronchitis in asthmatic children increase in association with long-term exposure to NO_2. Although health effects associated with NO_2 may be caused partially by other combustion-related pollutants that are emitted together with NO_2, the REVIHAAP study concluded that it is reasonable to infer that NO_2 has some direct effects as well (WHO, 2013).

Although UFP and BC are not regulated, there is increasing evidence of their association with health effects (WHO, 2013). For UFP, the evidence is still limited, but accumulating, and studies have shown that UFPs partly act through mechanisms not shared with larger particles that dominate mass-based metrics such as $PM_{2.5}$ or PM_{10} (WHO, 2013). Due to their small size, UFPs can be transported deep into the respiratory system allowing interactions with the lung tissue and potential absorption and translocation into the blood stream (Heal et al., 2012). Janssen et al. (2012) reviewed the studies on BC health effects and concluded that there is clear evidence of an association between both short-term (daily) variations in BC concentration and long-term BC exposure with health effects. The associations are even more robust than those with $PM_{2.5}$ or PM_{10}, suggesting that BC is a better indicator of harmful particulate substances from combustion sources (especially traffic) than undifferentiated PM mass (Janssen et al., 2011, 2012). McCreanor et al. (2007) showed the detrimental effect of exposure to pollution from diesel traffic, which includes UFP and BC, on the lung function in asthma patients. Further, in June 2012 the International Agency for Research on Cancer (IARC), an agency of WHO, declared diesel engine exhaust, of which BC is a major component, as a carcinogen (IARC, 2013). BC is not necessarily a major directly toxic component of fine PM, but it may operate as a universal carrier of a wide variety of chemicals of varying toxicity. BC is a valuable additional air quality indicator to evaluate the health risks of air pollution dominated by primary combustion particles, as well as benefits of traffic abatement measures (Janssen et al., 2011; Reche et al., 2011).

Environmental impact of air pollution

Air pollution not only has a health impact, it also affects the environment and the natural world. It contributes to eutrophication, atmospheric ozone, and the acidification of water and soil. It also impacts on agricultural production and forests, causing yield losses, and can damage materials and buildings (EEA, 2015b). Several air pollutants are also climate forcers. Tropospheric O_3 and BC are both

short-lived climate forcers that contribute directly to global warming. On the other hand, other PM components have a cooling effect on the climate (EEA, 2015a). BC plays an important role in the climate system because it absorbs solar radiation, influences cloud processes, and alters the melting of snow and ice cover. It is estimated to be the second most important human emission in terms of climate forcing, after CO2 (Bond et al., 2013). Because of it short lifetime, reducing BC emissions could slow the rate of global warming immediately (EPA, 2012; Bond et al., 2013).

2.1.3 Monitoring networks, policy and trends

In most industrial countries many of the above-mentioned pollutants are regulated and monitored in official monitoring stations. These monitoring networks consist of fixed monitoring sites where measurements are performed with well-established and comparable methods. The implementation of these networks is mainly driven by the need to comply with air quality standards (Kuhlbusch et al., 2014). European regulation defines the minimum number of monitoring locations for the different pollutants. This number depends on the measured concentration levels and the population density (EU, 2008). For example, for an agglomeration with a population between 250,000 and 500,000, a minimum of two stations is required if the maximum concentrations exceed the upper assessment threshold. The European monitoring networks also have to meet certain data quality objectives, and reference measurement methods are defined. Besides the data quality objectives for the fixed monitoring stations, less strict objectives are defined for indicative measurements. Those measurements can be used to supplement the minimum number of fixed monitoring stations (EU, 2008).

In Flanders, the Flemish Environmental Agency (VMM) monitors the air quality through a network of monitoring stations. The VMM has 61 monitoring stations spread across Flanders, but not every pollutant is measured at each station (VMM, 2015). The evolution of measured ambient concentration levels of NO_2, $PM_{2.5}$ and BC in Flanders is shown in Figure 2.3.

The fixed monitoring networks can characterize the temporal trends (e.g. Figure 2.3) and the regional variation well. However, the monitors that are used in the monitoring networks are mostly expensive (typically more than 10,000 EUR per component), and due to this high cost the spatial coverage of the monitoring networks is rather low (Theunis et al., 2016b). Therefore, those networks are not able to characterize the small-scale variation.

Table 2.1: Overview of the EU limit values, EPA air quality standards and WHO air quality guidelines for PM_{10}, $PM_{2.5}$, NO_2 and 03 (in $\mu g\,m^{-3}$). Note: for simplification of the table, no distinction is made between EU limit and target values. Depending on the pollutant, some limit values are allowed to be exceeded a certain number of times per year (EU, 2008; EPA, 2016b; WHO, 2006).

		Target ($\mu g\,m^{-3}$)		
Pollutant	Averaging period	EU	EPA	WHO
PM_{10}	1 day	50	150	50
	Calendar year	40	-	20
$PM_{2.5}$	1 day	-	35	25
	Calendar year	25	12	10
NO_2	1 hour	200	191	200
	Calendar year	40	100	40
O_3	Maximum 8-hour mean	120	140	100

Air quality guidelines

The European directive on ambient air quality and cleaner air for Europe (2008/50/EC) is one of the main policy instruments on air pollution in the EU (EU, 2008; EEA, 2015a). This directive formulates a strategy to reduce the harmful effects of air pollution on human health and the environment. Among other things, the directive defines target and limit values. An overview of the limit values for a selection of pollutants is given in Table 2.1. Further, the directive also includes limit values for sulphur oxide (SO_2), nitrogen monoxide (NO), lead (Pb), carbon monoxide (CO) and benzene.

The World Health Organization (WHO) defines target values as well, the Air Quality Guidelines (AQG). These guidelines are based on the available epidemiological research, and are more stringent than the EU limit values for most regulated pollutants. For example, the WHO formulates a limit of 10 $\mu g\,m^{-3}$ for the yearly average concentration of $PM_{2.5}$, while the limit value of the EU is 25 $\mu g\,m^{-3}$. The national ambient air quality standards (NAAQS) of the United States are added in Table 2.1 as well (EPA, 2016b).

Evolution of the air quality

The evolution of measured ambient concentration levels of NO_2, $PM_{2.5}$ and BC in Flanders is shown in Figure 2.3. For these pollutants, there has generally been a

decreasing trend over the last decade(s) (VMM, 2015). This decreasing trend can be observed in the rest of Europe as well (EEA, 2015a) and in most high-income countries (Brauer et al., 2016). The reduction in exposure to ambient air pollution contributed to significant and measurable improvements in life expectancy in the United States (Pope et al., 2009). Globally, in contrast, the population weighted mean concentration of $PM_{2.5}$ increased by 20.4 % between 1990 and 2013, mainly driven by trends in South Asia and China (Brauer et al., 2016).

Despite the improvements in air quality in Flanders and Europe, a large proportion of European population and ecosystems is still exposed to air pollution in exceedance of European standards and WHO Air Quality Guidelines (AQGs). For example, in 2013 9 % of the urban population in the EU-28 was exposed to $PM_{2.5}$ levels above the EU limit value and approximately 87 % was exposed to concentrations exceeding the stricter WHO AQG value for $PM_{2.5}$ (EEA, 2015a). Several studies also documented that adverse health effects of PM are observed at concentrations well below the current European limit values (e.g. Crouse et al., 2012; Beelen et al., 2014; Raaschou-Nielsen et al., 2013), highlighting the need for lowering the European air quality limit values (Harrison et al., 2014). Apte et al. (2015) argued that even modest improvements in $PM_{2.5}$ levels would result in surprisingly large avoided mortality, not only in heavily polluted areas in Asia, but also in Europe and North-America.

2.1.4 High spatial and temporal variability in the urban environment

The dynamics of primary emissions, dispersion and secondary formation induce important differences in pollution levels in space and time (Seinfeld and Pandis, 2006). But as these dynamics are pollutant-dependent, different air pollutants will also show different spatio-temporal patterns. Pollutants that are dominated by local source emissions will typically show a larger spatial variation. Especially in the urban environment, with a high density of local sources, and especially for traffic-related air pollutants such as NO_x, UFP and BC, these differences can occur on a small scale (Kaur et al., 2007; Cyrys et al., 2012; Peters et al., 2014; Wu et al., 2015). For example, Peters et al. (2014) measured differences in BC levels up to a factor of 5 on a scale of 50 to 200 m. $PM_{2.5}$ and PM_{10} typically show a smaller spatial variability compared to UFP and BC. This reflects the larger impact of the regional background concentrations on the $PM_{2.5}$ and PM_{10} levels due to long range transported secondary aerosol (Cyrys et al., 2012; Heal and Hammonds, 2014). But these background levels can show a large day-to-day variation (Peters

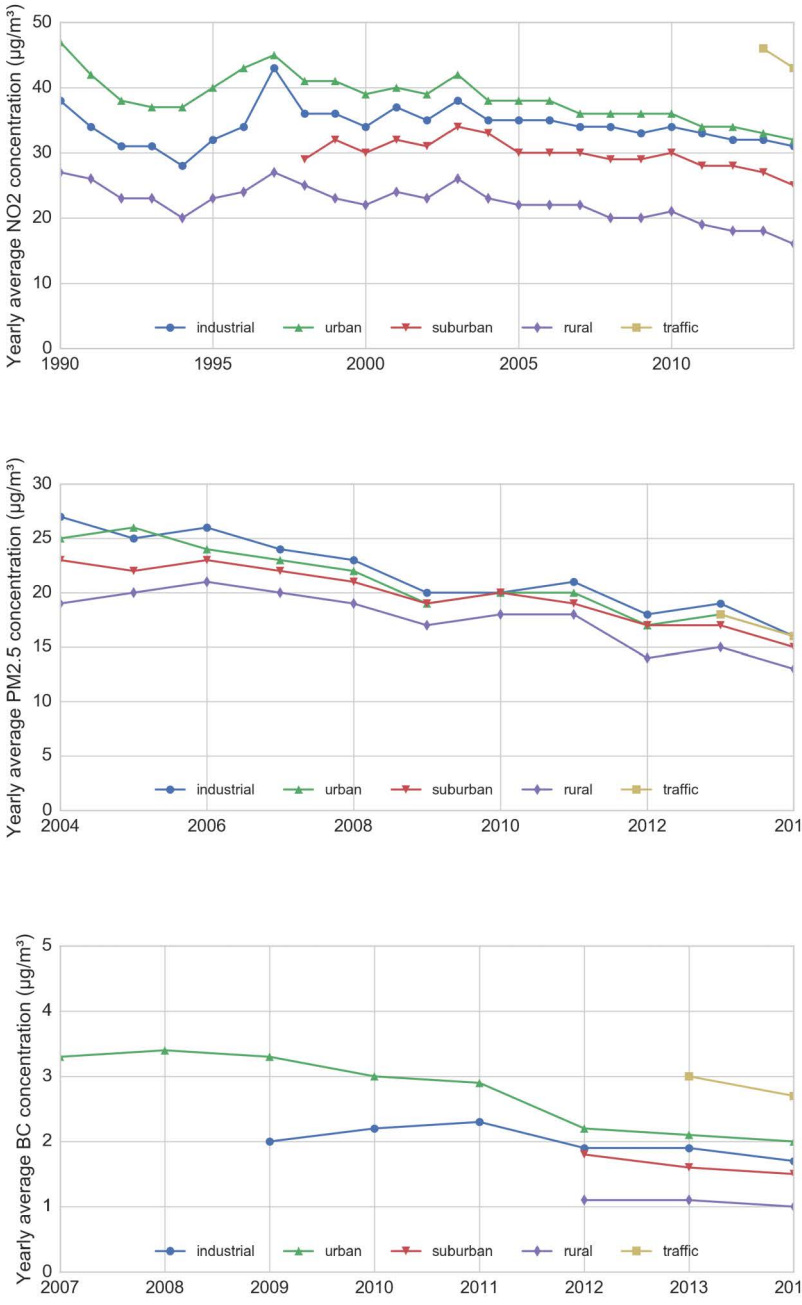


Figure 2.3: Evolution of the air pollution in Flanders: ambient concentrations for NO_2, $PM_{2.5}$ and BC based on the measurements at VMM stations for different environments (Source: http://www.milieurapport.be).

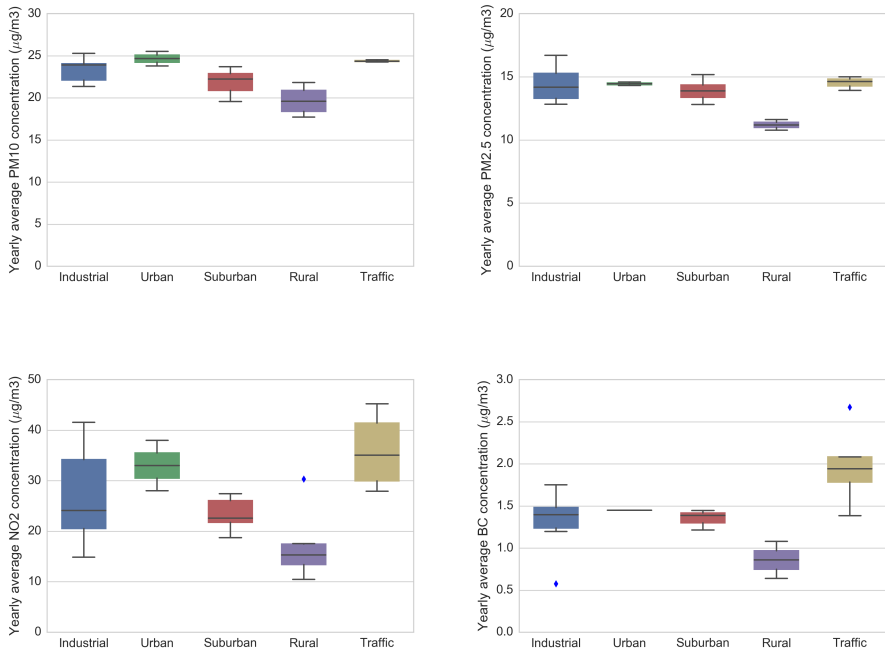


Figure 2.4: The average concentration for the year 2015 at the VMM monitoring stations in Flanders, grouped by the type of location. The figure is based on 56 stations for NO_2, 26 stations for $PM_{2.5}$ and PM_{10} and 17 stations for BC (Data source: www.irceline.be).

et al., 2013b; Wu et al., 2015). For UFPs, the regional background is less important because of its short atmospheric lifetime, which results in a high spatial variation (Klompmaker et al., 2015). The difference in spatial variability between pollutants in Flanders is illustrated in Figure 2.4, showing indeed a higher variability for NO_2 and BC compared to $PM_{2.5}$ and PM_{10}.

Within Europe, there are also substantial regional differences. In the ESCAPE study, air quality was measured at several study areas across Europe (36 for NO_2, 20 for PM, Cyrys et al. 2012; Eeftens et al. 2012b). Higher median NO_2 concentrations were found in Southern Europe, and higher PM_{10} concentrations in Southern and Eastern Europe. Regional differences could be due to, among other things, differences in traffic intensity, the fraction of diesel-powered vehicles, industrial emissions and climate. But, the variability within study areas is often larger than the regional differences. In case of NO_2, 40 % of the overall NO_2 variance was attributable to the variability between study areas and 60 % to variability within

study areas (Cyrys et al., 2012). The smaller-scale differences within the urban environment in both space and time and its drivers are discussed in the next two sections.

Temporal variability

The temporal variability in pollutant concentrations occurs at different time scales, and is mainly caused by changes in emissions or meteorological conditions. Dependent on the pollutant, typical seasonal or diurnal fluctuations can be noticed. But also large differences occur on a time-scale of hours to days, as well as on the very short-term time scale of seconds. Meteorological conditions (solar radiation, wind speed and direction, boundary layer height, temperature, humidity, rain, etc.) strongly affect the pollutant levels. It is difficult to relate the concentration levels to individual meteorological parameters, as the meteorological conditions are inherently linked to each other and can affect pollutants through direct physical mechanisms or indirectly through influences on other meteorological parameters (Pearce et al., 2011). In addition to meteorological conditions, changes in emissions are also an important factor. Traffic-related pollutants typically follow the diurnal pattern of traffic intensity, reaching a maximum during the morning rush hour, decreasing during the day because of atmospheric dilution processes, and increasing again in the evening (Reche et al., 2011). For example, UFP concentrations during the morning rush-hour period were found to be 2-5 times higher in comparison to night-time levels at traffic sites (dos Santos-Juusela et al., 2013; von Bismarck-Osten et al., 2013). On weekend days, this diurnal pattern changes because the morning rush hour disappears (Reche et al., 2011). Changes in emissions can also show seasonal patterns (e.g. residential heating in the winter).

The evolution of O_3 levels also shows a typical diurnal pattern. O_3 levels increase at midday coinciding with the maximum solar radiation (Reche et al., 2011). This effect also manifests at the seasonal level. Hot, dry summers will experience higher O_3 concentrations (Ramsey et al., 2014).

Wind speed and wind direction are two other important factors influencing the pollution levels. Wind direction determines the influx from upwind sources and influences the background levels. Low wind speeds, which can co-occur with temperature inversion and limited vertical mixing, will enhance the accumulation of pollutants near emission sources such as traffic in urban environments (Grundström et al., 2015). Higher wind speed, on the other hand, will reduce concentrations for most particulate and gaseous pollutants (Jones et al., 2010; Pearce et al., 2011; Grundström et al., 2015). But, it can also cause higher concentrations

in case of marine aerosol generation or wind induced resuspension from the land surface (Jones et al., 2010; von Bismarck-Osten et al., 2013).

Ambient meteorological conditions also influence the atmospheric transformation processes of aerosols (Hallquist et al., 2009; von Bismarck-Osten et al., 2013). For example, UFP concentrations are affected by temperature, causing differences between seasons (Padró-Martínez et al., 2012). Higher concentrations during winter months are attributed to higher UFP levels in vehicle exhaust and a greater atmospheric stability (Kumar et al., 2011; von Bismarck-Osten et al., 2013). Low temperatures favour the formation of ultrafine particles in the atmosphere by cooling and dilution of semi-volatile gases from vehicle exhausts (Charron and Harrison, 2003). Relative humidity will influence the chemical composition and size distribution of aerosol particles (Elminir, 2005; Sun et al., 2013)

Finally, at a small temporal scale (i.e. of the order of seconds), there can also be large fluctuations in the measured pollutant concentration due to emission events, for example the passing of a vehicle (Kaur et al., 2006).

Spatial variability

In the urban environment, large spatial differences are observed for pollutants that are dominated by local sources, typically traffic-related. Within the same city, large differences are found between traffic and background locations. In the ESCAPE study, ratios of the average concentration between street and urban background locations were calculated. For NO_2, this ratio was 1.63 on average for all sites over all study areas (ranging between 1.09 and 3.16 for the different study areas) (Cyrys et al., 2012) and for $PM_{2.5}$ absorbance, it was 1.38 (ranging between 1.02 and 1.77) (Eeftens et al., 2012b). dos Santos-Juusela et al. (2013) measured 5 times higher median PNC concentrations in a street canyon compared to an urban background.

The high spatial variability is in the first place caused by the local emission sources, but several other factors affecting the location-specific emission and dispersion characteristics will influence the actual pollution levels, resulting in a complex pollution pattern of urban micro-environments. The highest concentrations are typically observed along streets with high traffic volumes (Nikolova et al., 2011; Wu et al., 2015). But not only the traffic intensity is determining the emissions, also fuel type, fleet composition, driving conditions, speed, and road slope influence the actual emissions and consequently the concentration levels. For example, traffic intersections are pollution hotspots in the urban area because of the frequent stop-and-go situation causing changing traffic conditions with speed variations

and acceleration/deceleration. UFP emission can be up to an order of magnitude higher during acceleration compared with steady speed conditions (Goel and Kumar, 2014). Further, traffic does not only cause tailpipe emissions. Tyre abrasion and resuspension induced by traffic are also a major source of PM_{10} (and specifically of the coarse fraction between 2.5 and 10 μm) in urban areas, which can cause spatial differences for PM_{10} as well (Eeftens et al., 2012b).

Concentrations can rapidly decline with distance from the source (Kaur et al., 2007; Zhu et al., 2008; Hagler et al., 2010; Kumar et al., 2014). Baldwin et al. (2015) observed that concentrations halved at 89-129 m in the downwind direction and at 14-20 m upwind from the road. This concentration gradient depended on wind speed and direction. The distance between the receptor, e.g. a pedestrian or cyclist, and the traffic sources will therefore have a strong effect on the exposure. In a study of MacNaughton et al. (2014), bike lanes were found to have concentrations of BC and NO_2 that were approximately 33 % higher than bike paths, which are separated from the vehicle traffic. This was related to the higher distance from the road, but also to vegetation barriers and reduced intersection density.

Varying street geometry and building density introduce complex dispersion characteristics that further increase the spatial variability. A typical example is street canyons. The term street canyon refers to a relatively narrow street with buildings lined up along both sides (Vardoulakis et al., 2003). The geometry of the street canyon reduces the natural ventilation, creating hotspots of urban air pollution (Vardoulakis et al., 2011; Kumar et al., 2011; Weber et al., 2013; Peters et al., 2014; Wu et al., 2015). Even within streets, the spatial variability is very high due to the local dispersion characteristics (Vardoulakis et al., 2011). For example, it has been observed that the pollutant concentrations are higher on the leeward side of the street canyon (the upwind side in case of perpendicular roof-top wind flow) (Vardoulakis et al., 2003). As a result, the prevailing wind direction and the orientation of the street are also important factors. The presence of vegetation in street canyons can further limit the ventilation. This effect will often be larger than the pollutant removal capacity of vegetation, and in this way vegetation can rather lead to increased pollutant concentrations than to improved air quality, at least locally (Vos et al., 2013; Vranckx et al., 2015). At other locations, roadside vegetation barriers can also limit the exposure of receptors such as pedestrians (Al-Dabbous and Kumar, 2014).

2.2 Personal exposure to air pollution

Exposure is determined by the concentrations people experience in their living environments. When an individual moves from one place to another, his or her total personal exposure is determined by the time-weighted average of concentrations at each location (WHO, 1999; Vardoulakis et al., 2003). The actual dose is a product of exposure and dosimetry factors such as inhalation rate (Morawska et al., 2013). Small-scale spatial and temporal variability has a large influence on the actual pollution levels people are exposed to. Further, actual exposure is also highly dependent on individual time-activity patterns (Dons et al., 2011; Steinle et al., 2013). For example, in a personal exposure monitoring study of Dons et al. (2011), the exposure to BC differed considerably (by up to 30 %) for partners living at the same address. This difference could partially be explained by the time in transport.

2.2.1 Exposure in micro-environments

In most exposure studies, the concept of micro-environment is used to specify the locations that an individual visits (Steinle et al., 2013). In principle, a micro-environment is defined as a confined space where pollutant concentrations are assumed to be uniform, but given the high variability in pollution levels this assumption is often not true (Vardoulakis et al., 2003). In Belgium, most people spend around 71 % of their time at home, 10 % at work and approximately 6 % in transport (FPS Economy and Statistics Division, 2008; Dons et al., 2011). Typical micro-environments used in exposure studies are: indoor home, outdoor home, other indoor/outdoor (work, school) and transport (Steinle et al., 2013). In this thesis, the focus is on the transport micro-environment.

Transport micro-environment

Although not much time is spent in the transport micro-environment, it can be a significant contributor to the integrated daily exposure and to the variability in exposure between individuals, depending on the pollutant (Dons et al., 2011; Karanasiou et al., 2014). In the BC study of Dons et al. (2012), people spent on average 6 % of their time in transport, but it accounted for 21 % of the daily integrated exposure to BC and up to 30 % of the inhaled dose. The actual exposure during travelling shows a strong variability and depends on the route taken, the traffic intensity, duration and timing of the trip, vehicle specifications and fuel

type, and also transport mode (Zuurbier et al., 2010; Knibbs et al., 2011; Dons et al., 2013a).

The transport mode (private car, pedestrian, cyclist, several types of public transport) is an important factor for the exposure in transport. In literature, contrasting results can be found but in general lower concentrations are reported for cyclists or pedestrians than for car drivers or passengers inside vehicles (Kaur et al., 2007; Karanasiou et al., 2014). However, cyclists have two to five times higher respiration rates than travellers in motorized vehicles depending on the activity level (Int Panis et al., 2010; Bigazzi and Figliozzi, 2014). When taking that into account, the actual intake dose can be similar or even higher for the active modes (cyclists and pedestrians) compared to inside cars (McNabola et al., 2008; de Nazelle et al., 2012; Dons et al., 2012).

2.2.2 Exposure assessment

Assessing the exposure to air pollution at the personal or population level has three important applications: epidemiological studies, risk assessment and impact assessment of policies (WHO, 2005). For epidemiological studies, an estimate of exposure for each individual in the study is needed to investigate the links between disease and exposure. For risk and impact assessment, the effects on the health of the whole population or a sensitive subgroup that can be attributed to the exposure to a certain pollutant or to a change in exposure due to policy implementation will be estimated.

Traditionally, exposure assessment for epidemiological studies relies on pollutant concentrations from a limited number of fixed air quality network sites and static population distributions. Based on the population density or the home address of the study object, the exposure is estimated with annual average concentrations measured at nearby sites (Steinle et al., 2013; de Nazelle et al., 2013). However, those central monitoring stations may not be able to accurately characterize the spatial variability in the surrounding area and may thus not be representative for the whole city (Wilson et al., 2005; Vardoulakis et al., 2005; Santiago et al., 2013). Therefore, caution is advised when using central monitoring sites as proxies for exposure in epidemiological studies, as it could lead to substantial exposure misclassification for air pollutants showing high spatial variability (Wilson et al., 2005; Cyrys et al., 2012; Gu et al., 2013). In order to account for the within-city spatial variability in pollution levels, the concentration at the home address can also be derived from modelled concentration maps. Different methods to construct such pollution maps will be discussed in more detail in Chapter 9.

The traditional methods of exposure assessment based on the home address ignore the impact of the individual time-activity patterns (Beckx et al., 2009; Dons et al., 2011; de Nazelle et al., 2013; Steinle et al., 2013). The assumption that people are only exposed to pollution at their place of residence will lead to bias and loss of statistical power in health effects estimates (de Nazelle et al., 2013; Setton et al., 2011). For example, in the studies of Montagne et al. (2013, 2014), measured personal exposure was compared with an estimate of the residential exposure based on LUR models. The results varied between pollutants, but in general the LUR models did not predict measured personal exposure well within a single city.

Exposure estimates can be improved by determining more accurately where people spend their time (Panis, 2010; Setton et al., 2011). Alternative methods to reduce exposure misclassification are direct measurements (personal exposure monitoring) or indirect exposure models that take activity patterns into account.

Personal exposure monitoring

Personal monitoring provides the best estimate of the actual personal exposure (Mölter et al., 2012; Steinle et al., 2013; Dons et al., 2014b). By carrying a portable monitor, it enables to monitor in people's everyday environments instead of focusing on a single micro-environment (Steinle et al., 2013). Over the last years, more and more examples of personal exposure monitoring studies combining portable sensors with GPS have emerged (e.g. Dons et al., 2011; Broich et al., 2012; Dons et al., 2012; Steinle et al., 2015; Bekö et al., 2015; Ryan et al., 2015).

The development of small and light-weight devices made personal exposure monitoring possible. However, due to the relatively high cost and high burden on the participants, personal measurements are not feasible in many studies, especially not in epidemiological studies with a large number of individuals and over an extended period (Physick et al., 2011; Dons et al., 2014b; Su et al., 2015b). Furthermore, in some studies, exposure needs to be estimated retrospectively (Mölter et al., 2012). Therefore most current studies using personal exposure monitoring are limited to short periods in a small sample (Dons et al., 2014b). Further developments in sensing devices will possibly make personal monitoring feasible at a larger scale (Steinle et al., 2013; Snyder et al., 2013; de Nazelle et al., 2013). These developments will be discussed in Section 2.3.

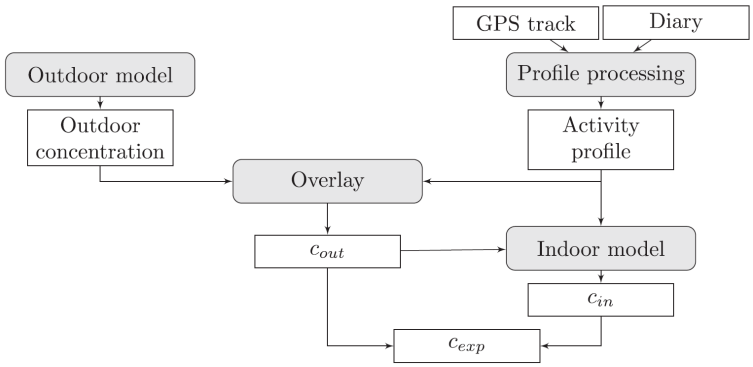


Figure 2.5: Overview of a personal exposure modelling methodology based on GPS tracks and diaries that are combined with a modelled pollution map (from Gerharz et al. 2013).

Modelling exposure

Exposure models are developed to estimate personal or population exposure that are not only based on the residential address and that avoid expensive personal measurements (Physick et al., 2011; Dons et al., 2014b). A number of models exist that combine the time spent in micro-environments with the estimated concentrations at each type of micro-environment (Gerharz et al., 2009; Dons, 2013). But these models are only based on the type of micro-environment and do not use the actual geographical position of the individual. Therefore, exposure models are developed that combine spatially explicit time-activity information and pollution models for different micro-environments. The time-activity information can be based on actual activity patterns (from GPS tracks and/or time-activity diaries) or based on simulated activity patterns (activity-based models, e.g. Beckx et al. 2009; Dons et al. 2014b).

Gerharz et al. (2009, 2013) developed a model for individual exposure estimation by combining existing models. GPS data and diaries provided the daily activity profiles and the associated geographical locations. This is then combined with the modelled pollutant concentrations in the different micro-environments for each time step. For the outdoor concentration levels, a dispersion model was used. Indoor concentrations were modelled separately based on the outdoor concentration and indoor activities estimated from the diaries. An overview of the approach is visualized in Figure 2.5. A comparison with personal monitoring data showed a good agreement and the model approximated exposure better than urban background station measurements (Gerharz et al., 2013). Similar models have been presented in other studies (e.g. Physick et al., 2011; Mölter et al., 2012; de Nazelle

et al., 2013). In the study of de Nazelle et al. (2013), the outdoor concentrations were adjusted based on ratios for different transport modes to model the transport micro-environment. Dekoninck et al. (2015) estimated BC exposure during cycling trips based on a city-wide noise assessment or noise map. The pollution maps that are developed in this thesis could be integrated in such a model chain for the transport micro-environment.

Tracking time-activity patterns

Steinle et al. (2013) identified a trend towards real-time tracking of individual time-activity patterns for personal exposure assessment. Traditionally, information on the time-activity patterns is recorded using a so-called time-activity diary (TAD), which can be laborious and the active cooperation of the participants is needed (Steinle et al., 2013; de Nazelle et al., 2013). This process can be facilitated by utilising GPS, and with the growing popularity of smartphones, which have integrated GPS systems, new opportunities emerge (de Nazelle et al., 2013; Su et al., 2015b). Using smartphone applications or the built-in tracking ability of smartphones, personal time-activity data can be collected more easily and potentially be used in large-scale exposure studies (de Nazelle et al., 2013; Glasgow et al., 2014; Su et al., 2015b).

2.3 From static monitoring to low cost sensing

2.3.1 The changing paradigm

Current advances in sensing technologies are leading to the development of low-cost, portable or miniaturised sensors that could be used as stand-alone devices, connected to smartphones or even embedded in smartphones. These new developments create new opportunities for detailed environmental monitoring and to involve communities and the general public in environmental monitoring (Aleixandre and Gerboles, 2012; Snyder et al., 2013; Piedrahita et al., 2014; Kumar et al., 2015; Theunis et al., 2016b). Portable sensors allow for mobile data collection, and low-cost sensors further allow the large-scale use of such sensors. These advances can supplement routine ambient air monitoring networks, provide citizens and communities with opportunities to monitor the local air quality and substantially enhance the ability to monitor personal exposure (Snyder et al., 2013). Snyder et al. (2013) talks about a changing paradigm (Figure 2.6). Historically,

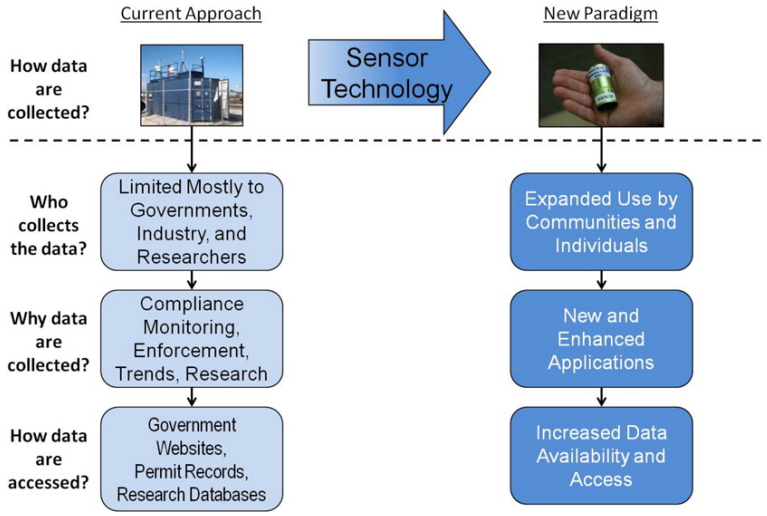


Figure 2.6: The changing paradigm in air quality monitoring (from Snyder et al., 2013).

approaches for monitoring air pollution generally use expensive, complex, stationary equipment, which limits who collects data, why data are collected, and how data are accessed. This paradigm is changing with the materialization of lower-cost, easy-to-use, portable monitors that provide high-time resolution data in near real-time.

Sensor devices are currently available to monitor a range of air pollutants, and new and improved devices are continually being introduced (Snyder et al., 2013; Castell et al., 2014). However, significant challenges remain with respect to the performance and use of the sensors. In the next two sections, a brief overview is given of the existing and emerging portable and/or low-cost sensors and sensing devices for monitoring both particle and gas concentrations.

2.3.2 Gas sensors

Part of the research activities during this PhD research was related to low-cost gas sensors (in the context of the EveryAware project). Low-cost sensors are commercially available for a broad range of gas species including NO_2, NO, CO and O_3. Prices range from a few to 30 EURO for metal oxide sensors, and from 50 to 80 EURO for electrochemical sensors. However, most of these sensors have not been designed to measure ambient air quality, but for industrial safety monitoring or emission testing. In these applications, typically much higher concentrations

are encountered compared to ambient levels (Theunis et al., 2016b). When these sensors are used for outdoor air quality measurements, the main issues are the inherent lack of sensitivity, sensitivity to changes in temperature and humidity, cross-interference by other gases, long-term stability and baseline drift.

Recently, Alphasense has a series of electrochemical sensors on offer that specifically target ambient air monitoring (e.g. O_3, NO_2, NO and CO) (Alphasense 2015), but specifications have still to be verified in real ambient conditions (Theunis et al., 2016b). Mead et al. (2013) presented a validation study of electrochemical NO, NO_2 and CO sensors (Alphasense, UK) that were optimised for ambient use. After applying data post-processing procedures to correct for baseline sensitivity to temperature and humidity and to correct for O_3 interference, they found a promising agreement with a co-located reference monitor. However, the validation was only conducted for a limited number of sensors and with a low time resolution (hourly averages). Hourly averages can be sufficient for stationary applications, but not for using the sensor in mobile monitoring, which was one of the applications presented in that paper.

Several commercial devices, integrating the basic sensors and strategies to improve the performance, are becoming available. This results in a higher complexity and significantly higher cost (from several hundred to several thousand EUR) of the final device (Theunis et al., 2016b). Examples of such devices are the AQmesh (Geotech) and the Aeroqual portable monitors. For example, Deville Cavellin et al. (2015) used the Aeroqual S500 in a measurement campaign to develop land use regression models. They noted large differences between individual Aeroqual devices, but good correlations with a co-located reference monitor. Until now, only few reports exist in which the use of these devices is compared to measurements from reference devices, generally reporting variable sensor performance under real-world conditions (Williams et al., 2014b; Jiao et al., 2015; Theunis et al., 2016b).

Sensor arrays and field calibration

A way to overcome the limitations of the low-cost gas sensors is the use of field calibration and the utilization of multivariate information from a set of different gas sensors and/or temperature and humidity sensors (sensor arrays) together with advanced regression techniques. Field calibration using co-located reference monitors can provide more accurate concentration estimates than laboratory calibration (Piedrahita et al., 2014). In sensor arrays, each of the sensors responds differently to the pollutant of concern and in this way it can provide a solution

for lack of selectivity of single sensors. Sensor arrays are also referred to as 'electronic nose'. De Vito et al. (2008, 2009) demonstrated the use of a sensor array for stationary outdoor measurements of benzene, CO and NO_2 at a traffic-intense location. More recently, Spinelle et al. (2015) compared the performance of several field calibration methods using an array of O_3, NO_2, NO, CO and CO2 sensors based on different working principles. The results for O_3 using a neural network were encouraging (within a limit of 30 % uncertainty), but for NO_2 still large values for bias and uncertainty were reported. Sensor arrays seem to have a high potential to counteract selectivity and calibration issues by combining different types of sensors. On the other hand, the field calibration requires a reference device to be deployed for a certain period next to the sensor array to develop the calibration model. The calibration model can be site and time specific, and the performance downgrades with time as the gas sensors deteriorate (Spinelle et al., 2015; Theunis et al., 2016b).

The EveryAware SensorBox

The EveryAware project investigated the use of off-the-shelf available low-cost sensors to estimate people's exposure to traffic-related pollution (Peters et al., 2013a). Part of the activities during this PhD research were carried out in the context of the EveryAware project. A portable sensor box with an array of electrochemical and metal oxide gas sensors was developed (Elen et al., 2012). Table 2.2 shows the result of a comparison of the individual sensor measurements with reference monitors between October 2012 and April 2013. The O_3 sensors showed a good correlation (0.83) with the reference O_3 measurements. Some CO sensors showed moderate correlation with the reference CO measurements, but a higher correlation with temperature. The correlations for the NO_2, NO_x and some of the CO sensors was low. The correlation between the NO_2 sensor and the O_3 reference was also high (0.64), indicating cross-sensitivity of the sensor.

Based on this sensor array, machine learning techniques were used for multivariate field calibration to estimate black carbon. The performance on independent data was, however, variable: R^2 between -0.37 and 0.59 and Pearson correlation between 0.33 and 0.77 for four different cities (Sîrbu et al., 2015). Due to the limited performance and usability for mobile monitoring of the sensor box, the work performed in this context is not further reported in this thesis.

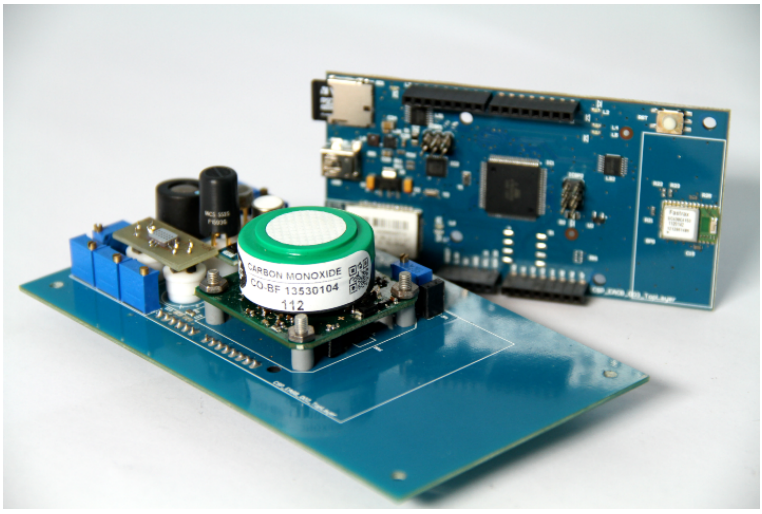


Figure 2.7: The EveryAware SensorBox: the two electronics boards with the gas sensors mounted on top of the sensor shield. These electronics boards were placed in a sensor chamber (box) with a pump to control the air flow (not shown).

Table 2.2: Correlation between 30 min averaged sensor measurements and co-located reference gas measurements from station BETR801 of the official Flemish air quality monitoring network and temperature (T) and relative humidity (RH) measurements from the sensor box. The average correlation of 4 sensors is shown together with the standard deviations between brackets.

	Reference monitors				T	RH
Sensors	CO	NO	NO_2	O_3		
Alphasense CO-BF	0.52 (0.16)	0.41 (0.11)	0.34 (0.11)	-0.32 (0.14)	-0.81 (0.11)	0.00 (0.16)
e2v MiCS-5521 CO	0.31 (0.04)	0.32 (0.04)	0.34 (0.04)	-0.09 (0.11)	0.89 (0.06)	-0.14 (0.06)
e2v MiCS-5525 CO	0.60 (0.02)	0.51 (0.05)	0.56 (0.05)	-0.71 (0.05)	0.50 (0.06)	0.25 (0.03)
Figaro TGS 2201 CO	0.25 (0.02)	0.32 (0.01)	0.17 (0.00)	-0.48 (0.01)	0.45 (0.03)	0.46 (0.07)
Figaro TGS 2201 NO_x	-0.78 (0.01)	-0.40 (0.06)	-0.24 (0.05)	0.47 (0.05)	-0.40 (0.04)	-0.21 (0.03)
e2v MiCS-2710 NO_2	-0.58 (0.02)	-0.40 (0.06)	-0.31 (0.08)	0.64 (0.07)	-0.40 (0.04)	-0.27 (0.02)
e2v MiCS-2610 O_3	-0.67 (0.06)	-0.56 (0.02)	-0.55 (0.05)	0.83 (0.07)	-0.18 (0.15)	-0.12 (0.19)
Applied Sensors VOC	0.63 (0.02)	0.43 (0.17)	0.53 (0.15)	-0.44 (0.26)	0.23 (0.22)	0.14 (0.10)

2.3.3 Particle sensors

Most existing particle monitors are still quite big, heavy and costly to be used in large scale applications, but lower-cost sensing devices for particles are becoming available (Theunis et al., 2016b). Examples are the Shinyei Technology optical particle counters and the Sharp optical dust sensor, both costing about 15 EUR. The Alphasense OPC-N1 optical particle counter is another small and light weight optical device for measuring particle concentrations of sizes between 0.5 to 15 μm aerodynamic diameter. Dylos corporation manufactures compact air quality monitors as well, and the latter two devices are more expensive (200 - 300 EUR).

Several monitoring case studies tested and used some of these devices (e.g. Holstius et al., 2014; Williams et al., 2014a; Steinle et al., 2015; Johnson et al., 2016). For example, Steinle et al. (2015) recently published a personal exposure study using the Dylos 1700 device. They concluded that the devices provide a robust representation of relative changes in particulate matter concentrations, but validation experiments need to be conducted over longer periods and at different locations. Johnson et al. (2016) tested three Shinyei particles sensors (PPD42NS, PPD20V, and PPD60PV) in a real-world mobile setting, but found varying performance in comparison to reference instruments (R^2 between 0.10 and 0.86). In a review of several low-cost PM sensors of different manufacturers by the EPA, Williams et al. (2014a) also concluded that the performance characteristics were highly variable between the devices. Most case studies using low-cost particle sensing devices are still in an experimental phase where a calibration function is developed based on simultaneous measurements with reference instrumentation. Long-term validity of these calibration functions and their transferability to other areas is largely unknown at this stage (Theunis et al., 2016b).

2.3.4 Challenges for low-cost sensing

We can conclude this section on low-cost sensing that, at this point in time, there are no readily available, low-cost solutions for ubiquitous outdoor air quality sensing (Theunis et al., 2016b). The performance of most of the available lower-cost sensors is not yet thoroughly characterized and their long-term reliability is unknown (Snyder et al., 2013). However, sensor technology is developing rapidly, and it is conceivable that in several years low-cost sensors will meet the current requirements of regulatory monitoring (EPA, 2013). Some encouraging examples show that the use of low-cost sensors has potential, but there remain a number of technical and practical challenges including the development of robust sensors that produce high quality data, careful electronics design, durability of the sensing

elements, advanced data post-processing or field calibration procedures, rigorous evaluations of sensors, and integration of data from multiple sensors of different quality (Snyder et al., 2013; Kumar et al., 2015; Theunis et al., 2016b)

Some portable, but more expensive instruments for measuring particles are already available on the market for quite some time. In absence of reliable low-cost sensors, this thesis will focus on one of these higher quality devices: the micro-aethalometer (microAeth AE51, AethLabs), a portable BC monitor. Other examples include the TSI DustTrak, which measures the mass concentration of different particle size classes between 1 and 10 μm, and the TSI P-trak, a condensation particle counter which measures the number concentration of particles smaller than 1 μm. The micro-aethalometer is smaller (easier to use as a portable device) and more user-friendly than the P-trak and DustTrak. Further, past experience has shown promising results with this instrument (e.g Dons et al., 2011; Van Poppel et al., 2013). An extensive validation of the micro-aethalometer is presented in Chapter 3. Although these devices are performing relatively well, they have to be compared on a regular basis with standard monitoring devices, are still rather expensive (roughly between 5,000 and 10,000 EUR) and need regular attention (e.g. replacing filters of the micro-aethalometer or refilling the P-trak). These elements limit their wide-spread use (Koehler and Peters, 2015; Theunis et al., 2016b), but a limited number of devices can already be used for mobile (participatory) measurements.

2.4 Citizen science and participatory monitoring

Citizen science encompasses a range of different ways in which citizens are involved in science (Conrad and Hilchey, 2011; SCU-UWE, 2013). This may include mass participation schemes in which citizens gather data, as well as smaller-scale activities such as local policy debates. Participatory monitoring and citizen science are often mentioned as ways to collect large datasets at a reasonable cost compared to classical data collection methods. But, citizen science can also be a way to use and rely on the knowledge and experience of people. It is recognized in many studies as a way to include stakeholders and the general public in the management of environmental issues (Conrad and Hilchey, 2011). Further, participatory monitoring or citizen science is also believed to raise the awareness and understanding of citizens (Snyder et al., 2013). Therefore, Theunis et al. (2016a) distinguish two objectives that should be considered in citizen science projects: the scientific and the social learning objective. The scientific objective focuses on the factual results of the scientific project, while the social learning objective is focused on processes

of creating shared knowledge and visions, awareness and behavioural change, co-creation and transition (Theunis et al., 2016a). While the concept of citizen science is not new, the number of projects has been growing due to the Internet and the use of smartphones and hand-held devices (Snyder et al., 2013). Some successful examples of citizen science include projects concerning biodiversity monitoring, volunteered geographic data, and personal weather stations.

Citizen science and participatory monitoring approaches can also be used in environmental monitoring, and more specifically air quality monitoring. Already in 2006, Burke et al. mentioned the theoretical potential of participatory sensing to investigate the relationship between air quality, urban traffic and public health. In the previous chapter, an overview was given of the emergence of lower-cost, easy-to-use portable air pollution monitors that provide high-time resolution data in near real-time. This provides opportunities for citizen science, and monitoring approaches based on low-cost sensors and smartphone apps are appearing in domains such as noise, air quality or radiation monitoring. Often the idea of pervasive or ubiquitous sensing is put forward, relying on a multitude of sensors with which data are collected in an uncoordinated almost effortless way, and on intelligent data post processing and mining (Kumar et al., 2015). However, this is not yet a reality today.

The potential for participatory environmental monitoring crucially depends on three strongly interdependent factors: the availability, quality and cost of monitoring tools (sensors, apps, etc.), sound data collection and data processing methods, and finally the participation of volunteers (Theunis et al., 2016a). The first factor, the monitoring tools, is currently still a limiting factor for wide-spread do-it-yourself approaches. But, some accurate portable instruments such as the micro-aethalometer are available, and those can already be used to set up participatory monitoring campaigns. But due to the high cost of these devices, such campaigns will likely have to be conducted in close collaboration with scientific institutions that can provide the instruments. The second factor, sound data collection and data processing methods, also increases the complexity of air quality research for citizen scientists. In the second part of this thesis, we focus on this factor by looking at mobile data collection and processing methods.

Over the last years, several citizen science or participatory monitoring projects have been set up in air quality applications. Some examples are given here. The EveryAware project developed a sensor box with an array of low-cost gas sensors and a smartphone application. As part of the project, an international competition was set up where citizens were involved in mobile air pollution monitoring using the sensor box, combined with a web-based game to monitor perceived levels of

pollution (Sîrbu et al., 2015). The CITI-SENSE project built a "citizens' observatory" observatory platform. This platform provides a toolkit for citizen scientists by integrating different sensing devices, smartphone apps, data processing and visualization techniques, and a web platform. In Antwerp, the AirBezen[1] project, led by the University of Antwerp, set up a participatory monitoring campaign by distributing strawberry plants. The plants were used to biomonitor the deposited PM. The project received a lot of interest of participants, and more than 1000 plants were distributed (which resulted in around 700 valid samples). Recently, Ringland (an environmental organization that focuses on the impact and the possible covering of the ring road of Antwerp) started the CuriezeNeuzen[2] campaign to measure NO_2. They want to deploy passive sampling tubes at 1000 locations across the city. US EPA also pays a lot of attention to citizen science. They included the support of citizen science projects in their roadmap for the next-generation air monitoring (EPA, 2013) and developed a citizen science toolbox to provide communities with the components needed to initiate monitoring studies (EPA, 2016a).

2.5 Conclusion

The urban environment shows a high variability in air pollutant concentrations on a small scale, especially for traffic-related air pollutants. New approaches in air quality monitoring enable to better characterize this variability, and they can provide more accurate exposure estimates. Portable monitors allow to perform mobile measurements and gather data at a high spatial resolution. Low-cost sensor allow for a larger-scale deployment and participatory monitoring.

Sensor technology is a quickly evolving field. However, currently, the available low-cost sensors cannot yet be readily used for ambient air quality monitoring. Validation studies indicate a variable performance. Further, using these sensors for mobile monitoring induces additional challenges regarding response time and the influence of the changing environment. Therefore, in this thesis, we will make use of a more expensive and accurate, but still portable instrument: the micro-aethalometer (and additionally the P-trak in Chapter 5).

[1] `https://www.uantwerpen.be/airbezen`
[2] `http://ringland.be/academie/curieuzeneuzen/`

CHAPTER 3

An extensive validation of the micro-aethalometer for personal exposure and mobile monitoring campaigns

3.1 Introduction

Black carbon (BC), a constituent of fine particles and a product of incomplete combustion, is associated with adverse health effects, including cardiovascular mortality and cardiopulmonary hospital admissions (Janssen et al., 2011, 2012). BC is not yet regulated, but would be a valuable additional air quality indicator to evaluate the health risks of air pollution dominated by primary combustion particles, as well as the benefits of traffic abatement measures (Janssen et al., 2011; Reche et al., 2011; WHO, 2013). The BC concentration levels show a high spatial and temporal variability, certainly in the urban environment (Wu et al., 2015). Due to this high variability, personal exposure to BC is strongly determined by time-activity patterns, and the exposure while in traffic can exceed exposure in other micro-environments (Dons et al., 2011, 2012).

A portable monitor can be used to characterize the BC concentration levels at a high spatial and temporal resolution and to monitor personal exposure. The micro-aethalometer (microAeth Model AE51, AethLabs) is a lightweight, pocket-sized monitor that measures BC in real time and at a high temporal resolution. The micro-aethalometer is increasingly used in several types of studies where rel-

atively short-term BC measurement campaigns are conducted. A non-exhaustive overview of studies is given below. For example, the micro-aethalometer is used for measuring exposure in studies on short-term acute health effects (e.g. Weichenthal et al., 2011; Louwies et al., 2015; Provost et al., 2016), for personal exposure monitoring (e.g. Delgado-Saborit, 2012; Dons et al., 2011, 2012; Buonanno et al., 2013; Nieuwenhuijsen et al., 2015) and for assessing exposure of cyclists (MacNaughton et al., 2014; Hankey and Marshall, 2015b), in vehicles (Apte et al., 2011), in the subway (Vilcassim et al., 2014) or during commuting in general (de Nazelle et al., 2012; Li et al., 2015). Further, the micro-aethalometer is used to characterize the temporal variation at near-road locations (e.g. Song et al., 2013; Liang et al., 2013) or to assess within-city spatial variation using short-term (Ruths et al., 2014; Klompmaker et al., 2015) or mobile measurements (Van Poppel et al., 2013; Peters et al., 2014; Brantley et al., 2014a; Van den Bossche et al., 2015; Wu et al., 2015; Van den Bossche et al., 2016), to measure the vertical profile with a balloon (Trompetter et al., 2013) and to characterize vehicular emissions (Ning et al., 2013; Tang et al., 2015; Kamboures et al., 2015). The micro-aethalometer is also used in modelling studies, mainly for building land use regression models (e.g. Saraswat et al., 2013; Dons et al., 2013b, 2014a; Dekoninck et al., 2013; Ghassoun et al., 2015; Hankey and Marshall, 2015a).

Despite this increasing use, the number of studies validating the performance of the micro-aethalometer is rather limited. Viana et al. (2015) compared six micro-aethalometers with a multi-angle absorption photometer (MAAP) during four periods of two to four days. Cheng and Lin (2013) evaluated the real-time performance of the micro-aethalometer at a traffic site by comparing it with the rack-mounted Aethalometer AE31, but used only one instrument. Cai et al. (2014) performed a more extensive evaluation for fixed site and personal exposure measurements using six instruments and also tested a diffusion drier inlet for reducing the effect of changes in temperature and relative humidity. Several issues affect the performance of the micro-aethalometer. The instrument is characterized by a high noise ratio, possibly limiting its use at a high temporal resolution. Further, a higher loading of the filter leads to an underestimation of the BC concentration.

In this chapter, the micro-aethalometer for use in personal exposure monitoring and mobile monitoring is validated. The validation is based on the use of the instrument in several campaigns over the last years. The micro-aethalometer is compared with other instruments, including the MAAP, and the use of a filter loading correction is evaluated. Several intercomparisons of multiple instruments are presented to investigate the bias and uncertainty of the micro-aethalometer.

Further, the use of the micro-aethalometer at a high temporal resolution is evaluated.

3.2 Materials and methods

3.2.1 Instrumentation

There exists no standard method for measuring BC. Moreover, there is currently no formal agreement on the terminology that considers all aspects of specific properties and measurement methods (Petzold et al., 2013). As a result, there is a lot of ambiguity in the scientific literature, with measurements that refer to BC using different terminology and are based on different properties of the particles. Different measurements methods cannot be easily compared. Therefore, Petzold et al. (2013) proposed more specific terminology that refers to the particular measurement method. Equivalent black carbon (EBC) should be used to denote black carbon measurements derived from optical absorption methods. Elemental carbon (EC) should be used for data derived from methods that are specific to the carbon content of carbonaceous matter. Refractory black carbon (rBC) should be used for measurements derived from incandescence methods. Black carbon (BC) can still be used as a qualitative description when referring to light-absorbing carbon-containing substances in atmospheric aerosol. However, for quantitative applications the term requires clarification on how the quantity is measured (Petzold et al., 2013). Frequently used instruments based on light absorption are the aethalometer (Hansen et al., 1984), the particle soot aborption photometer (PSAP, Bond et al. 1999) and the multi-angle absorption photometer (MAAP, Petzold and Schonlinner 2004).

The micro-aethalometer

The micro-aethalometer (microAeth Model AE51, AethLabs) is a lightweight, pocket-sized version of the aethalometer (Figure 3.1). This instrument measures the concentration of optically absorbing aerosol particles at a wavelength of 880 nm, and thus the reported values refer to equivalent black carbon (in $\mu g\,m^{-3}$), using a mass-specific absorption cross-section (MAC) of 12.47 $m^2 g^{-1}$. This MAC value is the default value used by the micro-aethalometer. MAC values vary in time and space depending upon source emissions, transformation during transport, etc. (Petzold et al., 2013). As a consequence, the MAC value, if measured,

Figure 3.1: The micro-aetholemeter (microAeth Model AE51, AethLabs).

would also vary when performing mobile measurements. However, this is a limitation of the (micro-)aethalometer, and more in general of filter-based absorption methods.

The concentration is derived from the rate of change in light attenuation of a particle spot on a filter through which air is pumped. The inlet flow rate can be set at 50-200 mL min^{-1} (for older version the maximum flow rate is 150 mL min^{-1}). A measurement timebase between 1 s and 5 min can be chosen. The response time is negligible. The optical measurement of the filter is instantaneous and the lag in time due to pumping of the air through the tube is less than a second for a flow rate of 150 mL min^{-1} and an inlet tube of 5 cm.

The multi-angle absorption photometer (MAAP)

The multi-angle absorption photometer (MAAP) is based on the simultaneous measurement of light transmitted through and scattered back from a particle-loaded filter at multiple detection angles, and is less influenced by aerosol light-scattering and filter loading effects (Petzold and Schonlinner, 2004; Petzold et al., 2005). Since the MAAP is used in the official monitoring stations in Flanders and several European countries to measure BC (EEA, 2013b), we decided to take this method as a reference to compare the micro-aethalometer with.

3.2.2 Measurement campaigns

This study makes use of several campaigns that were conducted over the past few years (some of which for this PhD research and others for third parties) and

that have already been published in research papers or reports. An overview of the campaigns and relevant references are given in Tables 3.1 and 3.3. The intercomparisons with multiple micro-aethalometer instruments all accompanied a field campaign, and were performed before and/or after the campaign for quality assurance. The measurements of the MAAP are obtained from reference monitoring stations of the Flemish Environmental Agency (VMM).

3.2.3 Data processing

Measuring BC with the micro-aethalometer at 1 s time resolution is challenging due to the occurrence of signal noise, especially at low concentrations. Hagler et al. (2011) proposed a method to reduce the noise while preserving the significant dynamic trends in the time series: the Optimized Noise-reduction Averaging (ONA) algorithm. This algorithm performs an adaptive time-averaging of the BC data, with the incremental light attenuation (ΔATN) through the instrument's internal filter determining the time window of averaging. It is based on a user-specified minimum change in attenuation ($\Delta\text{ATN}_{\text{min}}$) which determines the length of the averaging window. For a sufficiently high BC concentration and/or a long timebase, the ΔATN in each time step will be greater than $\Delta\text{ATN}_{\text{min}}$ and the time resolution will be preserved. However, for a low BC concentration and/or a short time base, ΔATN will be less than $\Delta\text{ATN}_{\text{min}}$ and the time series will be averaged over the time interval that is needed to reach an ATN increase of at least $\Delta\text{ATN}_{\text{min}}$. The average value is allocated to all time steps of the time interval. The choice of $\Delta\text{ATN}_{\text{min}}$ and the flow rate will determine the average length of the time window used in the averaging (higher flow rate gives a higher ATN increase per time step).

The effect of ONA is investigated for a high temporal resolution in Section 3.3.3. For the other sections, ONA processing is always applied on the raw data (with $\Delta\text{ATN}_{\text{min}}$ of 0.05).

3.2.4 Evaluation approaches

Based on the parallel measurements, metrics are calculated following the ISO20988 guidelines for estimating measurement uncertainty for air quality data.

Parallel measurements with a reference method For the comparison between the micro-aethalometer and the MAAP measurements, the following metrics

are used. Firstly, the bias $u_B(y)$ is the systematic error of the aethalometer compared to the MAAP calculated as the difference in the overall mean. Secondly, the standard uncertainty $u(y)$ of the measurement y (residual standard deviation) is expressed as the root mean squared error:

$$u(y) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(y(j) - y_R(j))^2}\,, \tag{3.1}$$

where $y_R(j)$ are the reference measurements for a series of $j = 1$ to N observations. Finally, also the R^2 of the linear regression of the micro-aethalometer on the MAAP is given.

Parallel measurements with identical measuring systems The bias for each instrument k is calculated as the mean of all measurements of that instrument minus the mean of all measurements of all instruments:

$$a(k) = \overline{y_k} - \overline{\overline{y}}\,, \tag{3.2}$$

with $\overline{y_k} = \frac{1}{N}\sum_{j=1}^{N} y(k,j)$ and $\overline{\overline{y}} = \frac{1}{K}\sum_{k=1}^{K}\overline{y_k}$. The overall bias is then calculated as:

$$u_B(y) = \sqrt{\frac{1}{K}\sum_{k=1}^{K} a^2(k)}\,. \tag{3.3}$$

To assess the uncertainty at a certain time resolution, the standard deviation in each time step is given as:

$$s(j) = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(y(k,j) - y_R(j))^2}\,, \tag{3.4}$$

where $y_R(j) = \frac{1}{K}\sum_{k=1}^{K} y(k,j)$ is the mean value of all instruments at time step j (regarded as the reference value in absence of a real reference). The standard uncertainty is then calculated as the root mean squared error of these uncertainties:

$$u(y) = \sqrt{\frac{1}{N}\sum_{j=1}^{N} s^2(j)}\,. \tag{3.5}$$

The instrument bias is the systematic deviation from the true value of an instrument averaged for all measurements. The overall bias gives an indication of the expected instrument bias. The standard uncertainty, on the other hand, gives the uncertainty of an individual measurement at one time step. It is expressed as the standard deviation of the population of possible results of a measurement

Table 3.1: Available datasets for the comparison with the MAAP. The average concentration level during the campaign is given with the interdecile range (10-90 % percentile) between brackets in $\mu g\,m^{-3}$.

Period	Length	VMM station	Micro-aeth	Average conc.	Reference
2011-01	44 days	44R801 (Borgerhout)	AE209	3.5 (1.0 - 7.0)	Dons et al. 2012
2012-02	25 days	44R802 (Borgerhout)	AE283	3.9 (1.0 - 7.3)	Peters et al. 2014
2014-03	14 days	44R750 (Zelzate)	AE347	2.2 (0.6 - 4.8)	Van Poppel, unpublished

(based on the readings of the different instruments at one time step). Given that the mean of all instruments is used as the the true value (in absence of reference measurements), the bias and uncertainty are an indication of the precision and not the accuracy of the micro-aethalometer.

3.3 Results and discussion

3.3.1 Comparison with the MAAP

In this section, a comparison exercise of the micro-aethalometer and the MAAP is presented. The goal is to compare the measured concentration levels and calculate the bias between the micro-aethalometer and the MAAP, to investigate the filter loading effect on the micro-aethalometer and to evaluate a correction algorithm. For the comparison, the MAAP is used as a 'reference' to evaluate the micro-aethalometer measurements.

Several datasets are available for the comparison exercise (Table 3.1). For each, one micro-aethalometer was positioned at a VMM monitoring station next to the inlet of a MAAP instrument. For each period, the MAAP data are available at half hourly resolution. The result for one campaign, the 2011 one, will be discussed in detail. For the other campaigns, an analogous analysis is performed and the results are included in the supplementary material (Appendix 3.A).

Results An overview of all available data of the 2011 campaign is given as a time series (Figure 3.2a) and scatter plot (Figure 3.2b). Comparing those data results in a R^2 of 0.84, a bias of 0.37 $\mu g\,m^{-3}$ and a standard uncertainty of 1.0 $\mu g\,m^{-3}$ on an overall mean of 3.5 $\mu g\,m^{-3}$. To investigate the effect of the filter loading on the measured concentration by the micro-aethalometer, the relative difference between the MAAP and micro-aethalometer is calculated for each time stamp i:

$$BC_{diff,i} = (BC_{MAAP,i} - BC_{aeth,i})/BC_{MAAP,i}\,. \tag{3.6}$$

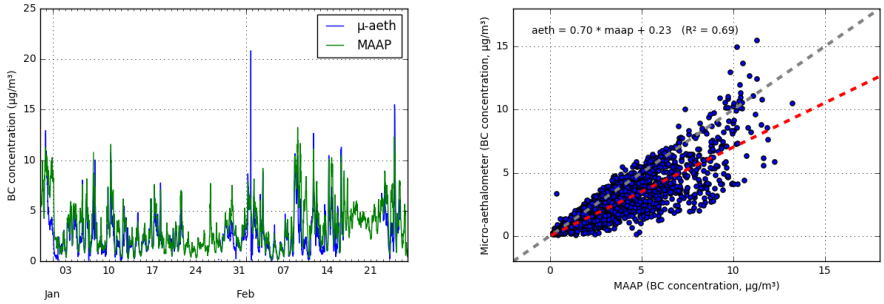


Figure 3.2: (a) Time series and (b) scatter plot of the micro-aethalometer and MAAP data for the 2011 data set (30 min time resolution). The red dotted line gives the linear regression of the micro-aethalometer on the MAAP (the grey dotted line gives the 1:1 line).

This relative difference $\mathrm{BC}_{\mathrm{diff},i}$ is plotted in function of the attenuation in Figure 3.3a. There is a clear positive relation with ATN. The micro-aethalometer underestimates the BC levels in comparison to the MAAP for higher ATN values. For an ATN value of around 100, this leads on average to a relative bias of 41 %. This systematic deviation in function of the ATN leads to a lower R^2 and higher mean bias. The relative differences between the micro-aethalometer and the MAAP are not dependent on the measured concentration (Figure 3.3b).

Filter loading effects were accounted for by a correction algorithm according to Virkkula et al. (2007):

$$\mathrm{BC}_{\mathrm{corrected}} = (1 + k \cdot \mathrm{ATN})\mathrm{BC}_{\mathrm{uncorrected}}\,, \tag{3.7}$$

where ATN is the attenuation. The compensation parameter k can be calculated for each filter spot so that the data become continuous over the spot change: complete Eq. (3.7) with the measurement of BC and ATN before ($\mathrm{BC}_{\mathrm{uncorrected}}$ and ATN) and after ($\mathrm{BC}_{\mathrm{corrected}}$) the filter change (Virkkula et al., 2007). This was done using the raw data at 5 min resolution and based on the average of the last and first 3 measurements before and after the filter change. Here, we used the average k value of all filters for the full dataset.

The filter loading correction is applied to three different cases: using all valid data, using all available data including the data with a 'high ATN' error message (status error code 4) and using only data with ATN < 75. The resulting metrics for these different cases are presented in Table 3.2. A value of k of 0.0076 was

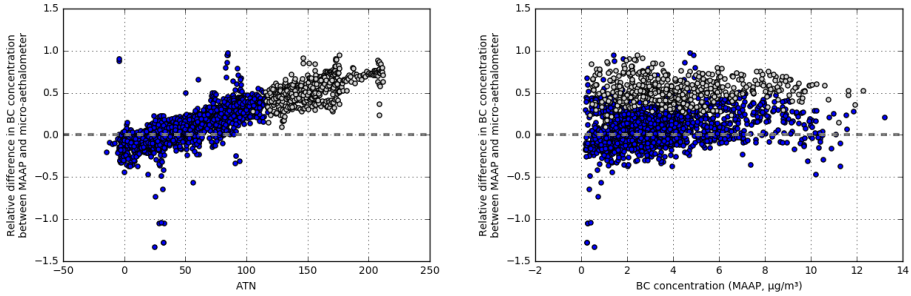


Figure 3.3: Influence of the attenuation on the measured concentration by the micro-aethalometer: the relative difference in BC concentrations between the MAAP and micro-aethalometer is plotted in function of the (a) ATN and (b) concentration measured by the MAAP. The data points where the instrument gave a 'high ATN' error message are displayed in grey.

obtained from the filter changes and used for the filter loading correction for the full dataset.

Based on the comparison with the MAAP, the micro-aethalometer shows a high standard uncertainty of 1.0 $\mu g\,m^{-3}$. This is partly due to the systematic bias with increasing ATN values. To investigate the effect of the filter loading correction on the standard uncertainty, both uncorrected and corrected data are adjusted for the bias by multiplying the values with the ratio of the overall mean of the MAAP and aethalometer measurements (resulting in a bias of 0). For the uncorrected data, this adjustment does not lead to a lower standard uncertainty. When applying both the filter loading correction and bias adjustment, the standard uncertainty is reduced to 0.77 $\mu g\,m^{-3}$ and the R^2 increases from 0.84 to 0.91 (Table 3.2). Note that the standard uncertainty is larger when only applying the filter loading correction in this case. This can be explained by the higher bias after filter loading correction which will also give a higher uncertainty (the loading effect masked a part of the bias). After applying the filter loading correction, the difference between the measured values by the MAAP and aethalometer instruments is no longer strongly related with the ATN value (Figure 3.4b).

The full dataset includes a lot of measurements with an error status code that indicates a high ATN value. When including these data in the analysis, the bias and uncertainty increase. Applying both filter loading correction and bias adjustment, the uncertainty is reduced strongly, but is still higher than with only the valid data. When adopting a stricter limit on the ATN value and only retaining

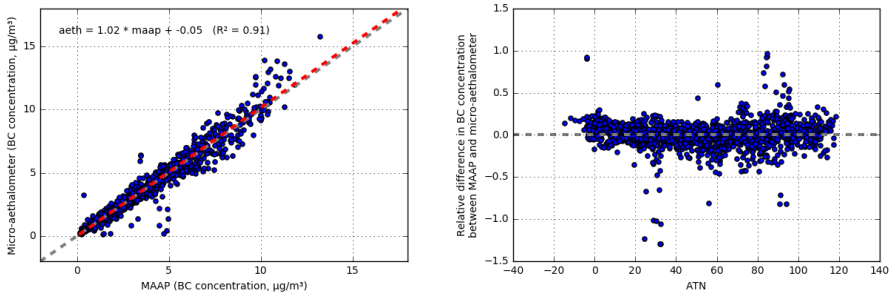


Figure 3.4: Comparison of the MAAP and micro-aethalometer after filter loading correction and bias adjustment: (a) scatter plot and (b) relative difference in function of ATN. The red dotted line gives the linear regression of the micro-aethalometer on the MAAP (the grey dotted line gives the 1:1 line).

Table 3.2: Results of the comparison between the MAAP and the micro-aethalometer for different cases: high ATN (all data including those with a 'high ATN' error message), valid data and low ATN (all data with ATN $<$ 75). Bias and standard uncertainty are given in $\mu g\,m^{-3}$.

	no correction	bias adjustment	filter loading	bias + filter
High ATN				
Overall bias	0.81	-	-0.59	-
Standard uncertainty	1.57	1.50	1.31	0.95
R^2	0.69	0.69	0.86	0.86
Valid data				
Overall bias	0.37	-	-0.78	-
Standard uncertainty	1.01	0.99	1.37	0.77
R^2	0.84	0.84	0.91	0.91
Low ATN				
Overall bias	0.06	-	-0.73	-
Standard uncertainty	0.54	0.54	1.09	0.41
R^2	0.94	0.94	0.97	0.97

the measurements with an ATN value below 75, a better result is obtained. The uncertainty without performing any correction (0.54 $\mu g\,m^{-3}$) is even smaller compared to using the valid data with correction for filter loading and adjustment for the bias. With the correction and adjustment, the standard uncertainty decreases slightly to 0.41 $\mu g\,m^{-3}$ and the R^2 is 0.97. This indicates that following the instrument's indication of a high ATN value is important for the data quality, and it is even better to change the filter ticket faster. Similar results are obtained for the other campaigns (Appendix 3.A).

The resulting bias between the micro-aethalometer and the MAAP for the valid data is -0.78 $\mu g\,m^{-3}$, which results in a relative bias of -22 %. Without filter loading correction, the bias is smaller (0.37 $\mu g\,m^{-3}$), but the underestimation of the BC concentrations at higher ATN values by the micro-aethalometer does mask the negative bias in this case. For the other two campaigns, a bias of -7 and +8 % is found.

Discussion The algorithm to post-process the aethalometer measurements for correcting filter loading effects (Eq. (3.7)) due to Virkkula et al. (2007) has some reported weaknesses. The algorithm was not designed to correct for all effects that influence the absorbance: it only corrects for the filter loading, and not for scattering effects (Virkkula et al., 2007; Collaud Coen et al., 2010). Collaud Coen et al. (2010) consider it as an unstable method and recommend not it to use it on atmospheric long-term datasets. Further, the value of the compensation parameter k depends on the type of aerosol, which induces a variation of k by season and location of measurement (Virkkula et al., 2007, 2015) and even by time of the day (Drinovec et al., 2015).

In this study, the compensation parameter k is determined following the same procedure as described in Virkkula et al. (2007). They determined the value of k based on filter changes: a value for k was calculated for each filter spot so that the data becomes continuous over the spot change. Virkkula et al. (2007) used the obtained value of k of each filter to correct the data obtained for that filter spot. In this study, however, we used the average k of all filters to reduce the noise in the obtained k values (due to a higher noise in the micro-aethalometer measurements compared to the instrument used by Virkkula et al. (2007)). This approach of determining the value of k based on filter changes requires relatively longer-term continuous measurements with the instrument, which is not always available in measurement campaigns. In Chapter 5, we took a different approach based on a shorter-term comparison with a MAAP but using the full time series (no continuous measurements including filter changes were available). In that case study, k

was estimated based on the comparison with the MAAP by fitting Eq. (3.7) using non-linear least squares (minimizing $BC_{\mathrm{MAAP}} - BC_{\mathrm{corrected}}$). In Dons et al. (2012), only those values with ATN < 75 were used, instead of applying a correction for the filter loading. This was also tested here both with and without filter loading correction. The aethalometer model AE33 (Magee Scientific) determines k continuously using a real-time loading effect compensation algorithm (based on measuring two filter spots with a different loading simultaneously) (Drinovec et al., 2015).

A good agreement between the micro-aethalometer and the MAAP was found (R^2 values of 0.91, 0.90 and 0.79 for the three campaigns). The bias between the micro-aethalometer and the stationary MAAP ranged between -22 and +8 % for the three campaigns. Similar results were obtained by Viana et al. (2015). Based on a comparison of six micro-aethalometers with one MAAP during four periods of two to four days, they found a relatively good agreement ($R^2 > 0.75$) and slopes ranging between 0.75 and 1.16). In the study of Cheng and Lin (2013), a micro-aethalometer was compared to an AE31 aethalometer and they observed that the BC levels measured by the micro-aethalometer were 14 % higher than those measured using the AE31 aethalometer. However, they only compared one micro-aethalometer instrument. They also observed that the BC levels were underestimated by up to 15 % when the ATN increases to 40, which is similar to what is found here.

To conclude, the lowest bias and uncertainty is obtained at low ATN values. It is therefore recommended to change the filter more often than the instrument indicates (an error message for high ATN values is generated by the instrument around ATN values of 100 to 120). How often this is depends on the flow rate and the measured concentration levels. For example, using a flow rate of 100 $\mathrm{mL\,min^{-1}}$, the filter reached an ATN value of 75 after 17 h of continuously measuring at an average concentration of 4 $\mu\mathrm{g\,m^{-3}}$. When high ATN values are not avoidable, the simple correction procedure according to Virkkula et al. (2007) can reduce the uncertainty caused by the filter loading when measurements are available to determine the value of k. To determine k, one needs to have a longer time series with multiple filter changes or a comparison with another instrument that corrects for filter loading effects.

3.3.2 Intercomparison of the micro-aethalometers

By means of an intercomparison exercise, the bias and uncertainty associated with the use of the micro-aethalometers is assessed. In this section, the micro-

aethalometer measurements are not compared to another method, but parallel measurements of several micro-aethalometer instruments are compared to assess the inter-instrument variability. As such, not the absolute bias in measuring BC is evaluated, but the relative bias compared to other instruments. This is relevant for campaigns using multiple instruments. Further, the standard uncertainty is evaluated for different time resolutions (5 and 30 min).

Intercomparisons of the instruments were performed by co-locating a set of micro-aethalometers for a period of time, which was typically done before and/or after a field campaign. Such measurements are available for several campaigns (see Table 3.3) at a 5 min time resolution. The measurements were not corrected for filter loading. For most of the campaigns, the ATN values were kept well below 75. Only for the MC3 and MC4 campaigns, higher values were reached, and all measurements with an error status were removed for the intercomparison exercise.

Results The results of one intercomparison campaign will be discussed in more detail (MC5_pre). The bias for each instrument ranges between -0.14 $\mu g\,m^{-3}$ and +0.16 $\mu g\,m^{-3}$, which results in an overall bias of 0.07 $\mu g\,m^{-3}$. Compared to the average concentration level during the campaign of 0.8 $\mu g\,m^{-3}$, this gives a bias of 8 %. The standard uncertainty is calculated for both a 5 min and a 30 min time resolution, and results in 0.10 $\mu g\,m^{-3}$ and 0.08 $\mu g\,m^{-3}$, respectively. The instrument bias is also illustrated in Figure 3.5 as the ratio of the instrument average to the average of all instruments. The ratios are given for the intercomparison before and after the campaign. For some micro-aethalometers, the sign of this ratio is consistent over the campaign, but for others it is not. When the measurements of the intercomparison after are scaled based on the ratios of the intercomparison before, the overall bias does not decrease. From this, it can be concluded that if the intercomparison shows a good correspondence between instruments ($< 10\,\%$), rescaling the micro-aethalometer measurements based on an intercomparison does not necessarily improve the measurements. The above analysis is repeated for the different campaigns, and a summary of the results is given in Table 3.3. The bias ranged between 1 and 12 %, and the standard uncertainty ranged mostly from 0.2 to 0.4 $\mu g\,m^{-3}$ at a 5 min time resolution and from 0.1 to 0.3 $\mu g\,m^{-3}$ at a 30 min time resolution.

Different flow rates were used in the campaigns (6 with 50 $mL\,min^{-1}$, 6 with 100 $mL\,min^{-1}$ and one with 150 $mL\,min^{-1}$). The campaigns with a lower flow rate have a slightly higher standard uncertainty (at a 5 min resolution): the median

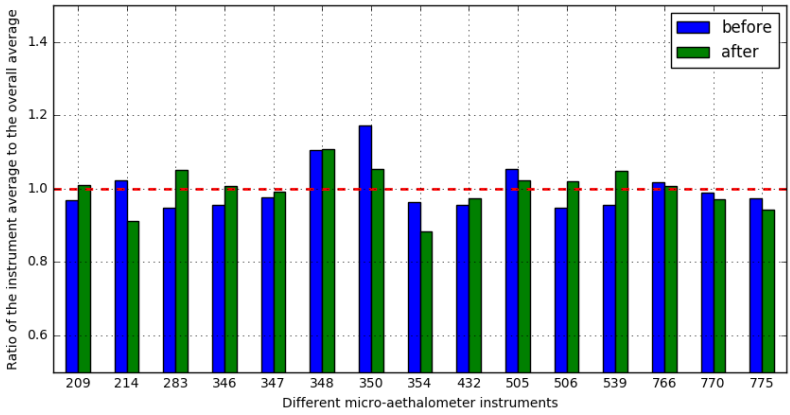


Figure 3.5: The relative instrument bias visualized as the ratio of the instrument average to the overall average for all instruments. The results of the intercomparison exercise before and after the MC5 campaign are given.

uncertainty for the different campaigns decreases from 0.32 to 0.24 $\mu g\,m^{-3}$ for 50 and 100 $mL\,min^{-1}$, respectively.

Discussion All instruments were serviced before each intercomparison before a measurement campaign including flow calibration, visual check of inlet tube and filter change. For the intercomparison after a measurement campaign, instruments were run without prior service. Some instruments showed a very large deviation from the other instruments. This can indicate flow errors, or the need for maintenance of the instrument (replacement of the pump, recalibration). For campaigns with multiple instruments, such an intercomparison exercise is therefore important. In the first place, the intercomparison can be used to identify broken or strongly deviating instruments in advance of the campaign. For the intercomparison exercises shown here, we left out on average 1 out of 10 micro-aethalometers because they deviated 25 to 50 % from the mean of all instruments. Our experience learns that such instruments can be used with some precaution, and regular intercomparison exercises are essential then to check the consistency in deviation of those instruments. In the results shown here (Table 3.3), only those instruments were kept that fell within 25 % of the average of all instruments.

It is clear that some bias between instruments can be expected when using multiple micro-aethalometer instruments. In the different intercomparisons, this bias ranged between 1 and 12 %. The intercomparison exercise can also be used to

Table 3.3: Results of the intercomparison of micro-aethalometers for the different campaigns. Some specifications of the campaigns are given: the duration of the intercomparison in hours, the number of instruments included, the average concentration level during the campaign with the inter-decile range (10-90 % percentile) between brackets. The concentration, bias and uncertainty are all expressed in $\mu g\,m^{-3}$.

	Period	Length (hours)	# aeths	Average conc (range)	Bias (rel.)	Standard uncertainty 5 min	30 min
MC1_pre	2011-11	21.9	11	3.0 (1.5 - 4.3)	0.29 (0.10)	0.46	0.35
MC1_post	2011-12	22.7	15	1.5 (0.9 - 2.2)	0.11 (0.07)	0.24	0.14
MC2	2012-03	20.2	6	3.2 (1.3 - 5.8)	0.23 (0.07)	0.34	0.30
MC3_pre	2012-09	89.2	4	3.6 (1.0 - 6.6)	0.03 (0.01)	0.25	0.11
MC4	2012-09	164.7	3	2.0 (0.5 - 4.0)	0.03 (0.01)	0.19	0.16
MC3_mid	2013-03	100.5	3	3.0 (1.1 - 5.6)	0.13 (0.04)	0.23	0.18
MC3_post	2013-07	143.5	4	2.0 (1.0 - 3.4)	0.04 (0.02)	0.34	0.28
MC5_pre	2014-01	46.3	18	0.8 (0.2 - 1.5)	0.07 (0.08)	0.10	0.08
MC5_post	2014-04	21.6	16	2.0 (0.9 - 3.0)	0.11 (0.06)	0.27	0.20
MC6	2014-07	24.6	19	0.7 (0.1 - 1.5)	0.08 (0.12)	0.80	0.40
MC7_pre	2014-11	29.6	11	2.3 (1.7 - 3.0)	0.19 (0.08)	0.31	0.27
MC7_post	2015-01	5.9	10	2.8 (2.0 - 3.8)	0.22 (0.08)	0.33	0.28
MC8	2015-05	53.3	4	1.0 (0.3 - 1.9)	0.12 (0.12)	0.34	0.21

rescale the different instruments to the overall mean, to reduce bias between instruments. However, in this case it did not lead to an improvement and the intercomparison of the unscaled measurements already showed good correspondence between instruments. For instruments that show a larger but consistent deviation, such a rescaling can be used. The bias between instruments also means that comparing absolute values of concentrations to other studies using micro-aethalometers should be done with care, as relatively large deviations are possible. For an intercomparison campaign it is important to have a dataset with representative variation in concentrations.

The uncertainty on the individual measurements is rather large, ranging mostly from 0.2 to 0.4 $\mu g\,m^{-3}$ at a 5 min time resolution. This decreases slightly when averaging the measurements to a 30 min time resolution (mostly 0.1 - 0.3 $\mu g\,m^{-3}$). The specifications of the microAeth(R) Model AE51 list a measurement precision of $\pm$ 0.1 $\mu g\,m^{-3}$ for 1 min average and flow rate of 150 $mL\,min^{-1}$. Using a higher flow rate generally leads to a lower standard uncertainty. Some studies equipped the micro-aethalometer with a higher volume sampling pump (800 and 1200 $mL\,min^{-1}$ in Weichenthal et al. (2011) and Ning et al. (2013), respectively) to reduce the uncertainty.

3.3.3 Use at a high temporal resolution

In the previous section, an intercomparison of multiple micro-aethalometers was performed using a temporal resolution of 5 and 30 min. However, for certain applications such as mobile monitoring, a higher time resolution than 5 min is needed. In this section, we will explore the performance of the micro-aethalometer at such a high time resolution. This analysis cannot be performed by comparing the micro-aethalometers to the reference MAAP device, because the measurements of the MAAP are not available at such a short time resolution (only the half-hourly values are stored by the reference monitoring station). Therefore, an intercomparison exercise is performed for sub-5 min resolutions and the impact of the processing of the measurements to deal with instrument noise is investigated.

For one of the campaigns used in Section 3.3.2, the MC2 campaign, data from the six micro-aethalometers are also available at 1 s resolution. Measuring BC with the micro-aethalometer at 1 s time resolution is challenging due to the occurrence of signal noise, especially at low concentrations. This leads to a high standard uncertainty and a lot of negative values in the raw data. Simply removing the negative values is inappropriate, as this would disregard the corresponding positive fluctuations due to noise and would lead to an overestimation of the concentrations (Hagler et al., 2011). Hagler et al. (2011) proposed a method to reduce the noise while preserving the significant dynamic trends in the time series: the Optimized Noise-reduction Averaging (ONA) algorithm (see Section 3.2.3 for more details on the algorithm).

For evaluating the use of the micro-aethalometer with measurement time base lower than 5 min, the standard uncertainty will again be calculated. It is calculated for different time resolutions between 1 s and 5 min, and using different preprocessing methods: the raw data, using the ONA algorithm with different values for $\Delta\mathrm{ATN}_{\mathrm{min}}$ and using a moving average with an averaging window of 5 and 30 s. Besides the standard uncertainty, also the mean correlation (mean of the Pearson correlations of one of the instruments with all others) and the number of negative values are calculated for the different scenarios.

Results The correspondence between the different micro-aethalometers at a high time resolution is illustrated in Figure 3.6a. Figure 3.6b shows the different preprocessing methods for a single instrument. In the previous section, we obtained an overall bias of 0.23 $\mu\mathrm{g\,m^{-3}}$ and a standard uncertainty of 0.34 and 0.30 $\mu\mathrm{g\,m^{-3}}$ for 5 and 30 min averages, respectively (Table 3.3). The standard uncertainty and mean correlation is calculated for different time resolutions between 1 s and 5 min

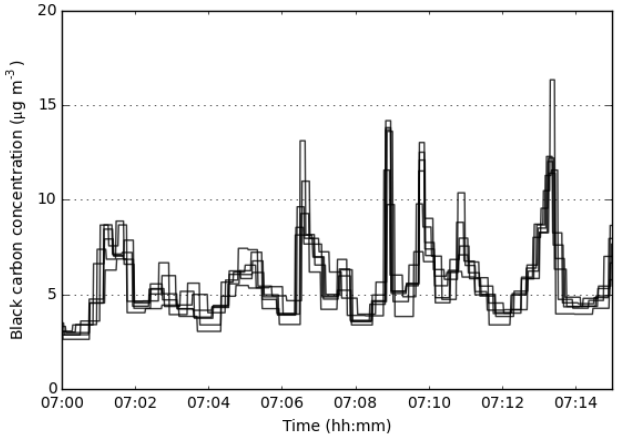
and using the original measurements as well as processed data using ONA and a moving average (Figure 3.7). Using the raw 1 s data leads to a high standard uncertainty above 3 $\mu g\,m^{-3}$ due to the large short-term (second-to-second) instrumental noise. Combined with the average concentration of 3.2 $\mu g\,m^{-3}$ during the campaign, this leads to a large number of negative data (16 %). By decreasing the time resolution, the standard uncertainty also decreases. When going from 1 s to 5 s time resolution, the standard uncertainty decreases from 3.2 to 1.0 $\mu g\,m^{-3}$, indicating that already a large part of the noise is reduced by using a 5 s time resolution, but a small percentage of negative values remained.

With the ONA-processed data, the 1 s time resolution already gives a good result (standard uncertainty of 0.8 $\mu g\,m^{-3}$, correlation of 0.93). However, although the data still have a 1 s time resolution, they are the result of averaging most of the data over a larger period (median averaging window of 45 seconds), so a better result is to be expected. The median averaging window varied in time, and during the rush hour it decreased to 20 seconds. Using the moving average also gives a good improvement in uncertainty and correlation compared to the raw data, but this method does not fully eliminate the negative values (2.7 % negative values at 1 s resolution).

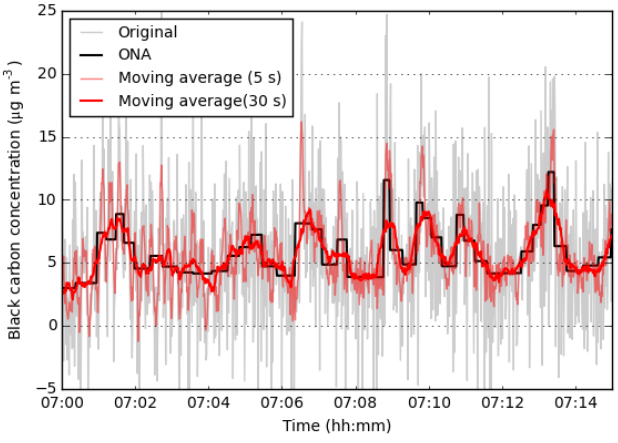
A value of 0.05 for ΔATN_{min} was used in the analysis above. Different values were tested and 0.05 generally gave the best standard uncertainty and correlation. Decreasing ΔATN_{min} to 0.01 decreased the median averaging window from 45 to 10 seconds, retaining more of the short-term dynamics, while the standard uncertainty increased only slightly from 0.8 to 0.9 $\mu g\,m^{-3}$.

Discussion Based on the standard uncertainty and correlation, there is not much difference between using the ONA algorithm and a moving average. Also, for lower time resolutions (> 15 s), the original measurements give almost the same performance. However, the ONA algorithm has the advantage of eliminating the negative values, while keeping the high time resolution where this is desired (periods with large dynamics in the concentration levels). When using larger time bases, the differences become smaller. A value for ΔATN_{min} can be chosen in the order of magnitude of 0.01 to 0.05, depending on the average concentration levels, instrument flow rate and the trade-off between preserving short-term dynamics and low standard uncertainty.

Considering the median averaging window of 45 s, the value of collecting data at a 1 s resolution and the suitability of the micro-aethalometer for mobile measurements can be questioned. However, the above analysis is based on stationary measurements and the actual averaging window of the processed data will vary

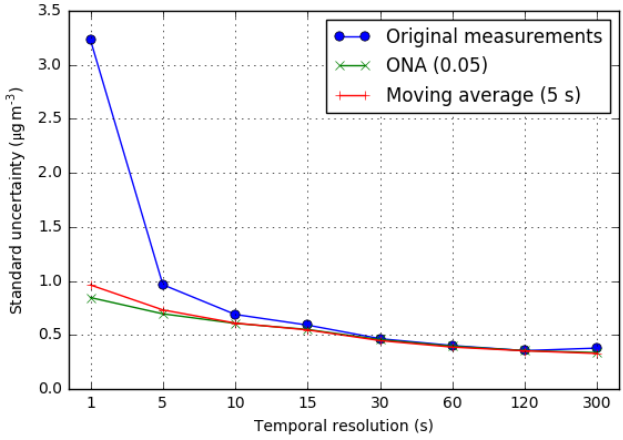


(a)

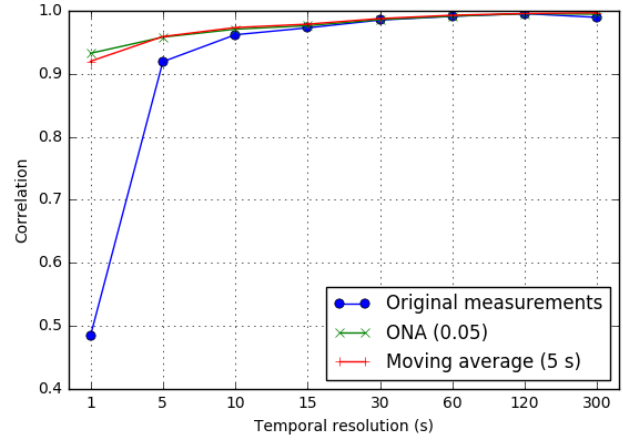


(b)

Figure 3.6: Illustrative part of the time series (15 min) of the full campaign: (a) The ONA-processed measurements of the six instruments and (b) the raw data and processed data using different methods for one of the instruments.



(a)



(b)

Figure 3.7: The (a) standard uncertainty and (b) mean correlation for different time resolutions and for different preprocessing methods. The mean correlation is calculated as the mean of the Pearson correlations of one of the instruments with all others.

from time to time and from place to place. It is expected that in mobile applications the median value will be lower at certain locations (as there can be higher concentration levels depending on the location). For example, in the mobile targeted campaign that is discussed in Chapter 5 and 6, the median averaging window during the mobile measurements was 15 s. This median averaging window further decreased to 12 s during the rush hour (between 8 and 9 am). Given that those are median values, half of the measurements will have a lower averaging window. Further, due to the algorithm, it are exactly those parts of the measurement series where high temporal resolution is necessary (e.g. high concentration peaks) that will have a lower averaging window in the ONA averaged data. Therefore, it is justified to collect data at a 1 s resolution as this enables an optimal resolution of the processed data.

After processing, a good agreement between instruments can be observed, even at a high temporal resolution. When using a time resolution of 1 to 10 s, the correlation is above 0.9, although the instruments have a higher standard uncertainty (0.6-0.8 $\mu g\,m^{-3}$) compared to a 5 min time resolution (0.3 $\mu g\,m^{-3}$). These results show the potential of the micro-aethalometer to observe high temporal dynamics in BC concentration levels. It is therefore, provided that the raw data are processed, a suitable instrument to perform mobile measurements.

3.4 Conclusions

This chapter evaluated the performance of the micro-aethalometer, a portable and easy-to-use BC monitor suited for personal exposure monitoring or mobile measurements. Some recommendations regarding the use of the micro-aethalometer can be put forward. First, a regular comparison with another BC monitor or with other micro-aethalometers is needed to identify broken or strongly deviating instruments. When conducting a campaign with several instruments, the intercomparison is also needed to assess the bias between instruments, as relatively large deviations are possible. Second, it is recommended to change the filter ticket more often than indicated by the instrument status to prevent high filter loading and underestimation of the BC concentration. Third, the micro-aethalometer can be used at a high temporal resolution when using a higher flow rate and post-processing the data to reduce the standard uncertainty. However, a higher flow rate also results in a higher noise level, which can hinder the person carrying the device in participatory studies.

Some aspects of the performance of the micro-aethalometer could not be evaluated with the available data sets. For example, it is known that the power source or changes in the temperature and humidity (e.g. when moving from indoors to outdoors) influence the quality of the measurements. To further investigate those issues, it could be useful to set up a tailored monitoring campaign.

This chapter confirms the potential of the micro-aethalometer as an instrument to perform mobile measurements. It shows a good agreement with the MAAP for relatively longer-term measurements, making it suitable for personal exposure monitoring. The micro-aethalometer is also able to observe high temporal dynamics in BC concentration levels, confirming its use to perform mobile measurements at a high temporal resolution.

3.A Comparison with the MAAP: results of other campaigns

Table 3.S1: Results of the comparison between the MAAP and the micro-aethalometer for the 2012 Borgerhout campaign (Peters et al., 2014). See Table 3.2 for more explanation.

	no correction	bias adjustment	filter loading	bias + filter
High ATN				
Overall bias	0.60	-	-0.16	-
Standard uncertainty	1.38	1.34	0.95	0.91
R^2	0.74	0.74	0.87	0.87
Valid data				
Overall bias	0.36	-	-0.28	-
Standard uncertainty	1.04	1.03	0.89	0.78
R^2	0.84	0.84	0.90	0.90
Low ATN				
Overall bias	0.11	-	-0.38	-
Standard uncertainty	0.87	0.87	0.93	0.75
R^2	0.87	0.87	0.91	0.91

Table 3.S2: Results of the comparison between the MAAP and the micro-aethalometer for the 2014 Zelzate campaign (Van Poppel, unpublished). See Table 3.2 for more explanation.

	no correction	bias adjustment	filter loading	bias + filter
High ATN				
Overall bias	0.43	-	0.21	-
Standard uncertainty	1.12	1.09	0.96	1.00
R^2	0.74	0.74	0.79	0.79
Valid data				
Overall bias	0.36	-	0.17	-
Standard uncertainty	0.98	0.98	0.90	0.94
R^2	0.76	0.76	0.79	0.79
Low ATN				
Overall bias	0.21	-	0.08	-
Standard uncertainty	0.76	0.76	0.72	0.74
R^2	0.82	0.82	0.84	0.84

PART II

MOBILE MONITORING STRATEGIES

CHAPTER 4

Introduction to mobile monitoring in air quality applications

4.1 Introduction

With the onset of portable monitors, mobile monitoring is increasingly used to acquire air quality data at a high spatial and temporal resolution in the complex urban environment. Mobile monitoring refers to the collection of data while moving along a path. As such, mobile measurements provide a solution for short-term studies to obtain a spatially spread data set that would not be feasible through stationary measurements, but at the expense of temporal representativeness.

Studies that use a mobile platform to measure at multiple locations for a certain time period more easily are sometimes also referred to as mobile monitoring studies. For example, stationary measurements can be performed using a mobile laboratory van (e.g. Vardoulakis et al., 2005; Pirjola et al., 2014) or portable instruments can be used by stopping for a short time period at predefined sampling locations (e.g. Merbitz et al., 2012; Klompmaker et al., 2015; Ghassoun et al., 2015). However, in this thesis we will focus on monitoring continuously with a mobile platform, without stopping purposely.

4.2 Applications of mobile monitoring

Mobile air pollution monitoring has several applications, but two main tracks can be distinguished: personal exposure monitoring and systematically mapping an

area. The first group of studies applies personal monitoring by directly equipping the study subjects with portable monitors. For example, this type of mobile measurements is used to assess the personal exposure of individuals during their daily activities (e.g. Dons et al., 2012; Ryan et al., 2015; Bekö et al., 2015) or to specifically assess exposure during trips with different modes of transport (e.g. Kaur et al., 2005; Int Panis et al., 2010; Dons et al., 2013a; Kingham et al., 2013)

Mobile monitoring can also be used to systematically map an area or route to construct air pollution maps at a high spatial resolution. Mobile measurements are performed with different platforms, e.g. pedestrians (Zwack et al., 2011a; Sabaliauskas et al., 2015), bicycles (Peters et al., 2013b; Sullivan and Pryor, 2014; Peters et al., 2014; Hankey and Marshall, 2015b), trams (Hagemann et al., 2014; Hasenfratz et al., 2015) and cars (Westerdahl et al., 2005; Hu et al., 2012; Hudda et al., 2014; Patton et al., 2014b). Several objectives can be pursued by systematically mapping an area. For example, mobile monitoring is used to study the spatial and temporal variation in air pollution (Weijers et al., 2004; Zwack et al., 2011c; Peters et al., 2014; MacNaughton et al., 2014), to identify pollution hot spots and relating them to specific sources, such as highways (Padró-Martínez et al., 2012; Choi et al., 2012; Patton et al., 2014b) or an airport (Hudda et al., 2014; Hsu et al., 2014), or to develop and validate air quality models (Zwack et al., 2011a; Patton et al., 2014a; Hankey and Marshall, 2015a; Dekoninck et al., 2015; Weichenthal et al., 2016b).

Further, mobile monitoring is also used for other purposes, including source apportionment (Argyropoulos et al., 2016), measuring emission factors in real driving conditions (Ježek et al., 2015), or studying spatio-temporal correlation with noise (Weber, 2009; Dekoninck et al., 2013). In the remainder of this chapter, and in this thesis in general, we will focus on the application of mobile monitoring to systematically map street-level air pollution.

4.3 Implications of the temporal variability

The high temporal variability in pollution levels, especially for pollutants with important local sources such as traffic-related pollutants, has already been discussed in Chapter 2. This variability and its cause differ from pollutant to pollutant, but we can distinguish different levels of temporal variability relevant for the context of mobile monitoring: day-to-day, within-day and micro-scale variability. Day-to-day variability is mainly caused by the meteorological conditions and urban

background fluctuations. Within-day variability is additionally related to the traffic dynamics, with significant peak concentrations at the morning and evening rush hours. Finally, at a small temporal scale (i.e. of the order of seconds), there are also large fluctuations in the measured pollutant concentration due to emission events, typically the passing of a vehicle. Mobile measurements, which can be seen as time-space series of data, result in only a short snapshot at a certain location in time and the measured values cannot necessarily be generalized as a typical value for that location.

To summarize, mobile monitoring allows to increase the spatial coverage of measurements, but at the expense of temporal representativeness. This leads to some critical issues that will be investigated in the next chapters. When mobile monitoring is used to derive high resolution maps that are representative for the typical concentration levels, large amounts of data are required to represent the range of possible meteorological and traffic conditions. Repeated measurements at all locations will have to be performed. Those repeated measurements can be obtained using different monitoring strategies.

4.4 Mobile monitoring strategies: targeted versus opportunistic

Mobile data collection can be performed in a targeted or an opportunistic way. We define **targeted mobile monitoring** as a coordinated, goal-driven approach in which the mobile measurements are deliberately planned and carried out with a specific purpose in mind. Efforts are typically concentrated in a well-defined area (e.g. a number of streets) over a specific time frame (e.g. during the morning peak hours) in an attempt to get a representative picture of reality. Literature examples of targeted data collection approaches using fixed routes are Hagler et al. (2010), Hsu et al. (2014), Pattinson et al. (2014), Peters et al. (2013b), Peters et al. (2014). This contrasts with **opportunistic mobile monitoring**, which we define as data collection making use of existing carriers to move measurement devices around. The movement of the carriers (the travelled route) is uncontrollable from the point of view of the researcher, as it is not designed and performed with the data collection in mind as primary goal. The data collection takes advantage of existing mobile infrastructure or people's common daily routines. This can be city wardens, parking wardens, public transport vehicles, taxis, street cleaners, bike couriers or postmen, but as well commuters that cycle every day to work. The participants do not decide on measurement location and time out of their interest

to monitor a given event. They do not envisage to cover a specific period of time, nor a specific location or route. Opportunistic mobile monitoring is a promising approach to collect large data sets that give useful additional information at a reasonable cost compared to classical data collection methods. But, depending on the set-up of the data collection, these data can lead to new challenges in data processing and interpretation.

4.4.1 Opportunistic mobile monitoring

The concept of opportunistic data collection as we have defined it above is closely related to opportunistic people-centric sensing as presented by Campbell et al. (2008). They describe a new sensing paradigm leveraging humans as part of the sensing infrastructure. By using small computational devices carried by individuals in their daily activities, information related to human activity and to the environment around them can be sensed opportunistically (Campbell et al., 2008; Kapadia et al., 2009; Kumar et al., 2015). As the data originates from sensors carried by people, new challenges for information security and privacy have to be addressed (Kapadia et al., 2009). However, in this study, we do not restrict it to people and focus more on the uncontrolled aspect of the data collection and its implications for data processing and interpretation of the results.

Opportunistic data collection can take different forms. We can consider two axes along which opportunistic data collection campaigns can vary. Firstly, they can vary according to the degree of human interaction they need. Possible human interactions are related to carrying the measurement system, the operation and maintenance of the measurement system and to the data collection and handling. Campaigns based on sensors mounted on vehicles such as cars, buses or trams can run independently for long periods without human interaction after initial set-up. The more human interaction the data collection needs, the more the user-friendliness of the measurement system and the motivation of the people involved become important issues. Secondly, the data collection can be structured or unstructured. Structured data collection is performed by repeatedly following fixed routes, possibly within a fixed time frame.

The studies of Hasenfratz et al. (2015) and Hagemann et al. (2014) are examples of opportunistic campaigns with a structured data collection, as they performed mobile measurements with sensors installed on the roof of public transport vehicles (trams). In these cases, the measurements are restricted to the route of the bus or the tram tracks. Another example of fixed routes are commuters performing mobile measurements (e.g. Weichenthal et al., 2008), although in this case the

route taken is more flexible. Aoki et al. (2009) built an environmental air quality sensing system and deployed it on street sweeping vehicles. The street sweepers most likely still follow rather fixed routes, but they typically follow different routes on different days leading to a more unstructured dataset. The case study described in Chapter 7 with city wardens carrying measurement devices during their daily surveillance tours results in unstructured measurements without distinct patterns in space or time. The studies of Hasenfratz et al. (2015) and Hagemann et al. (2014) require limited human interaction as the sensor nodes are supplied with power from the vehicles and data are transmitted automatically (Hasenfratz et al., 2015). When using humans as the mobile carrier (Weichenthal et al., 2008, this thesis), more human interaction is needed (carrying, switching on/off, charging the device, etc.). Other possibilities to gather data opportunistically would be to work with postmen, taxi drivers, parking wardens, and so on.

Opportunistic data collection is not restricted to the air quality examples above. Noise mapping using smartphone apps (e.g. NoiseTube, D'Hondt et al. 2013) can also be done in an opportunistic way. Floating car data coming from so-called probe vehicles (i.e. vehicles that are equipped with the necessary devices to transmit data to a data centre at regular time intervals) are used to map the congestion and speed. These probe vehicles can be deployed specifically, but this information can also be retrieved from existing uncontrolled cars equipped with a GPS. The term 'opportunistic data collection' is also used in another context within the field of Wireless Sensor Networks (WSN). Here, opportunistic data collection is the transfer of the data from static, disconnected sensor nodes to a central base station by mobile 'mules' that pass by both the sensor node and the base station (Tseng et al., 2013). In this way, the sensor nodes are opportunistically connected with their central servers. This application deals with data transfer and is not about performing measurements, and as such is distinct enough not to cause confusion about terminology.

The characteristics of opportunistic monitoring as outlined above hold some consequences for the processing and interpretation of the data. From the perspective of the researcher, there is no control over the specific location and time of the measurements. This can result in sampling bias where certain urban microenvironments are underrepresented or absent in the data. For example, a postman or a parking warden will not frequently enter green zones such as parks. Similarly, there can also be a bias due to the time of the measurements. In the case of data collection by commuters, the measurements are mainly limited to the rush hours. The sampling can also be biased by the weather conditions, e.g. when the data collection stops when it rains, or when the commuter takes the car instead of the

bicycle on rainy days. As a consequence, there will be different measurement conditions for different locations, hindering the comparability of the results between these locations. This is a major problem, as it complicates the data interpretation, making it less evident to use the results for air quality mapping.

4.4.2 Participatory mobile monitoring

Mobile monitoring also plays an important role in the context of the growing interest in participatory air quality monitoring and citizen science. Targeted monitoring campaigns can be planned and/or conducted with the involvement of a limited number of volunteers. This will, however, require highly committed participants who are prepared to spend time and efforts to repeatedly cover the targeted monitoring area on the appropriate moments (Theunis et al., 2016a). When low-cost sensors with sufficient performance become available, data can be collected on a larger scale in an un-coordinated, opportunistic way. Participatory monitoring is a possible way to collect the large datasets needed for obtaining representative maps from mobile monitoring at a reasonable cost compared to classical data collection methods.

4.5 Conclusion

Mobile monitoring is increasingly used for mapping air quality. It can also play an important role in participatory monitoring campaigns. However, several issues on spatial and temporal representativeness can interfere with the real-life applicability of this type of monitoring. Given the temporal variability in pollutant concentrations, a limited number of mobile measurements may not be representative and may have limited value in assessing the air quality at a specific location. Therefore, suitable monitoring strategies with a sufficient number of repeated measurements and a careful processing and interpretation of the results are required. Data coming from unstructured opportunistic measurements may exhibit some additional challenges, as this approach often leads to data that are unevenly spread in space and time.

These topics are investigated in the next chapters. An extensive, targeted mobile monitoring campaign was conducted. The dataset is presented and the variability in air pollution is characterized in Chapter 5. Because of the size of the dataset in this case study, it allows to address the research questions about obtaining a representative and consistent image of the spatial variability of the urban air quality.

These methodological issues are investigated in Chapter 6. Then, an unstructured opportunistic campaign is presented in Chapter 7, and the potential of the opportunistic monitoring strategy to assess the spatial variability is explored.

CHAPTER 5

Cyclist exposure to UFP and BC in the urban environment

5.1 Introduction

An extensive, targeted mobile monitoring case study was set up to characterize the exposure of cyclists to ultrafine particles (UFP) and black carbon (BC). UFP and BC are both strongly traffic-related pollutants and show a high variability in space and time in the urban environment. The correlation between UFP and BC in the traffic micro-environment has been investigated in several studies. Westerdahl et al. (2005) report a high correlation ($R^2 = 0.76$) at a 60 s time resolution based on mobile measurements in the Los Angeles area. In another study based on mobile measurements in a provincial town in Belgium (Van Poppel et al., 2013) the correlation (R^2) between UFP and BC was 0.41, based on averages per traffic zone for individual runs. Hagler et al. (2009) report a high correlation ($R^2 = 0.65$) between 10-minute averaged UFP and BC measurements for a near road sampling campaign at a major roadway in Raleigh, North Carolina.

The objective of this chapter is to examine UFP and BC pollution data from mobile monitoring at different scales. A multi-scale approach is taken to investigate (i) the spatio-temporal variability of UFP and BC concentration in the urban environment, (ii) to study the correlation between UFP and BC and (iii) to assess the potential of mobile monitoring for exposure assessment of a cyclist. The temporal detail ranges between weeks, day and hour of the day. Spatially, the data

This chapter is based on the following publication: Peters, J., Van den Bossche, J., Reggente, M., Van Poppel, M., De Baets, B. and Theunis, J. (2014). Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. *Atmospheric Environment*, 92 31-43.

are treated at the suburban scale, the street level and up to a 10 m resolution. Although this study is confined in space and time, results could be generalized to other, potentially more extensive, cases.

5.2 Material and methods

5.2.1 Mobile monitoring in an urban environment

An extensive monitoring campaign using a mobile monitoring platform, the Aeroflex (Elen et al., 2013), was set up in Antwerp (51°12' N, 4°26' E, medium-sized city of 480,000 inhabitants, 985 inhabitants km^2), Belgium, during 4 weeks in February-March 2012. The Aeroflex is a bicycle equipped with compact air quality measurement devices to monitor ultrafine particle number concentration and black carbon concentration at a high temporal resolution (up to one second). Each measurement is automatically linked to its geographical location and time of acquisition using GPS (GlobalSat BU-353[1]) and internet time. Two air pollutants were monitored in this study: ultrafine particles (UFP, measured as particle number concentration (PNC), $\mathrm{pt\,cm^{-3}}$) and black carbon (BC in $\mu\mathrm{g\,m^{-3}}$). A TSI P-Trak ultrafine particle counter (model number 8525) was used to measure the number concentration of UFP (size range 0.02 to 1 μm) within a range from 0 to $5 \cdot 10^5$ $\mathrm{pt\,cm^{-3}}$ at a temporal resolution of one second. The P-Trak is based on the condensation particle counting technique using isopropyl alcohol (TSI, 2013). The flow rate of the P-Trak was 0.7 $\mathrm{L\,minute^{-1}}$, the temperature during operation was within the limits stated by the manufacturer, and alcohol saturation of the wicks was guaranteed. BC measurements were done with a micro-aethalometer (MicroAeth Model AE51, Aethlabs, see Chapter 3 for a detailed description of the instrument). The filter ticket was changed at the start of each measurement day, the inlet flow rate was set at 150 $\mathrm{mL\,min^{-1}}$ and measurements were made at a temporal resolution of one second. Measuring BC with the micro-aethalometer at one second time resolution is challenging due to the occurrence of signal noise, especially at low concentrations. Therefore, a noise reduction algorithm was used (Hagler et al., 2011, see Section 5.2.3).

A controlled monitoring set-up with fixed routes and time slots was used. Two fixed routes were defined before the start of the monitoring campaign. Route 1 was an approximately 2 km long loop at the Southern side of a busy arterial road (Plantin en Moretuslei), route 2 was a 5 km long loop at the Northern side of this road (Figure 5.1). The start and stop location for both routes was a reference

[1] `http://www.canadagps.com/BU-353spec.html`

air quality monitoring station from the Flemish Environmental Agency (VMM, station BETR801). Preference was given to use a controlled monitoring set-up with fixed routes and time slots to guarantee sufficient temporal and spatial coverage. Repeated measurements in time were preferred over a larger spatial coverage that could have been gained from monitoring along longer or random routes. In the same area and season, a monitoring campaign was performed in 2010 at four fixed locations, showing a spatial and temporal pattern of UFP (Mishra et al., 2012).

The major part of the routes was located in residential and commercial areas. Route 1 passed by a low-traffic square (Dageraadplaats), whereas an urban green background area (Stadspark) was crossed by route 2. Streets of different configuration and with different traffic dynamics were included in this study (Table 5.1). The traffic data used in this study are modelling results from the Traffic Centre Flanders based on traffic counts. The air quality monitoring campaign consisted of 256 runs along route 1 and 96 runs along route 2 spread over 11 days in the period between 2012-02-13 and 2012-03-08. The majority of the runs occurred between 7 a.m. and 1 p.m. on weekdays (Figure 5.2), covering the morning rush. Preference was given to have many repetitions in time for a part of the day over a sparser dataset that would have been obtained by making measurement runs at random hours. For organizational reasons, a measurement day was limited to 6 hours. A total of around 248,000 data points were collected for route 1 and 152,000 for route 2.

Table 5.1: Description of some street characteristics of a selection of streets along monitoring route 1 (R1) and route 2 (R2).

	Street name	Abbrev.	Description	Speed limit (km/h)	Configuration		Traffic density[a] (vehicles per day)	Length[b] (m)
					Nr lanes	Biking lane		
R1	Plantin en Moretuslei	PM	Entrance road	70	2	yes	43,381	890
	Kleinebeerstraat	KB	Residential	50	1	no	1,269	105
	Lange Altaarstraat	LA	Residential	50	1	no	5,585	120
	Wolfstraat	WO	Residential	30	1	no	5,680	145
	Dageraadplaats	DP	Public square, very low traffic	/	/	no	NA (∼0)	170
R2	Plantin en Moretuslei	PM	Entrance road	70	2	yes	43,381	995
	Stadspark	SP	Urban green	/	/	yes	NA (∼0)	565
	Quellinstraat	QU	Tertiary	50	1	no	9,590	315
	F. Rooseveltplaats	FR	Square, bus stops	50		partly	21,422	70
	Carnotstraat	CA	Entrance road	50	1	yes	22,963	430
	Provinciestraat	PR	Tertiary (street canyon)	50	1	no	12,174	690
	Bleekhofstraat	BL	Residential	30, 50	1	no	4081	490

[a] Modelled traffic density from the Traffic Centre Flanders based on traffic count data (includes both light and heavy traffic).

[b] Length of part of the street that is included in the route.

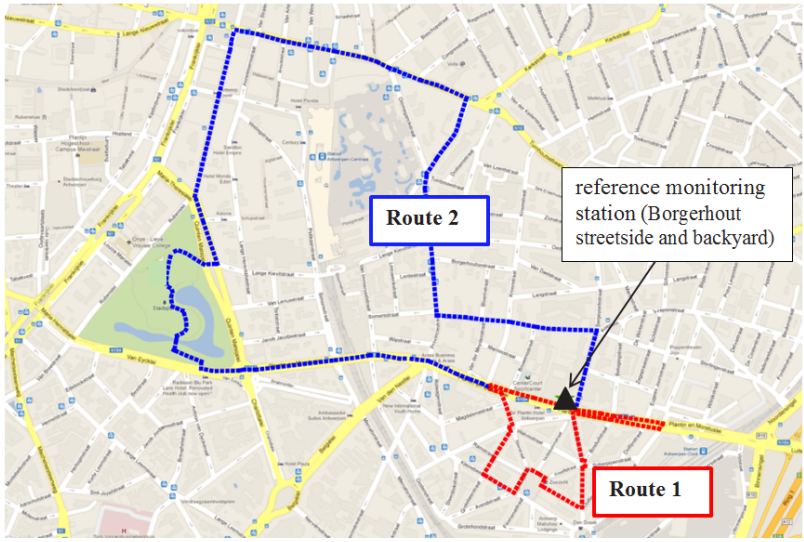


Figure 5.1: Overview of the bicycle routes: route 1 (red) and route 2 (blue).

5.2.2 Meteorological conditions

Meteorological data were obtained from a station of the Flemish Environmental Agency (VMM) near the study site (approx. 12 km). Wind conditions are measured at a height of 30 m, temperature and humidity are measured at 3 m. An overview of the temperature, humidity and wind conditions for February-March 2012 is given in Suppl. 5.S1. Averaged meteorological conditions during the hours of mobile monitoring for each day of the measurement campaign are given in Table 5.2. Temperatures ranged between 2.1 and 8.6 °C and the relative humidity was moderate to high (between 71 and 95 %). Wind of low to moderate speed from NE, NW and SW directions were most prevalent.

5.2.3 Data validation and quality assurance

In addition to the mobile monitoring, stationary measurements were made at a station of the AQ monitoring network at Plantin en Moretuslei (Figure 5.1) for data quality control. The individual P-traks and micro-aethalometers were put together for in total approx. 20 hours at the AQ monitoring station to identify possible bias between instruments. The micro-aethalometers were compared to a MAAP (Multi-Angle Absorption Photometry (Petzold and Schonlinner, 2004; Petzold et al., 2005)) instrument from the monitoring station, while the P-Traks were only compared with each other. Overall the three P-Traks showed a good

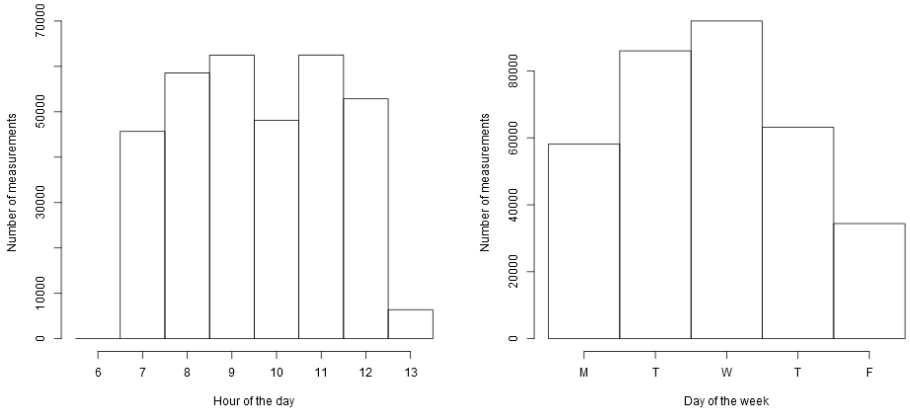


Figure 5.2: Overview of the number of measurements made in function of (a) hour of the day and (b) day of the week

Table 5.2: Overview of the meteorological conditions given as average temperature, relative humidity (RH) and wind speed and median wind direction during the monitoring hours of each of the monitoring days.

Date	Meteorological conditions			
	Temp (°C)	RH (%)	Wind speed (m/s)	Wind dir[a] (degrees)
feb/13	2.1	95	1.7	22
feb/14	4.1	89	3.5	85
feb/15	6.1	84	5.6	59
feb/20	1.9	74	2.4	327
feb/21	4.8	71	4.0	337
feb/22	5.5	71	4.3	326
feb/29	8.6	95	2.3	348
3/jan	7.8	95	1.1	250
3/feb	8.6	90	3.7	205
3/jun	5.2	80	2.8	251
3/aug	5.7	77	2.9	48

[a] 360° North, 90° East, 180° South, 270° West.

correlation (average R^2 of 0.97) and a slope ranging between 0.9 and 1.1. Therefore we did not rescale the P-Trak measurements.

The three micro-aethalometers showed a good agreement after noise reduction. Noise reduction was realised using the Optimized Noise-reduction Averaging algorithm (ONA algorithm, Hagler et al., 2011), with an attenuation threshold of 0.05. Rescaling parameters were derived from the comparison of the individual micro-aethalometers with the MAAP instrument from the monitoring station. Before determining the rescaling parameters, filter loading effects were accounted for by an additional correction algorithm according to Eq. (3.7) as explained in Chapter 3 (Virkkula et al., 2007). The parameter k was estimated by fitting Eq. (3.7) using non-linear least squares, for each micro-aethalometer independently. As the values did not differ significantly, an average value for k of 0.009 for all micro-aethalometers was used. For the fitting, the MAAP data were used as $\mathrm{BC}_{\mathrm{corrected}}$, as this device is considered to be less influenced by the loading effect compared with the micro-aethalometer (Petzold et al., 2005). Based on 30 min means, regressions between the corrected BC concentrations of the micro-aethalometers and the MAAP with slopes between 0.83 and 1.13 (R^2 of 0.97-0.98) were obtained. These slopes were used together with Eq. (3.7) to correct the mobile micro-aethalometer data. After the correction, the micro-aethalometer data showed a mean absolute error of 0.3 $\mu\mathrm{g\,m^{-3}}$ compared to the MAAP based on 30 min means (leading to a relative error percentage of 7 % for the calibration period). When looking at the agreement between the micro-aethalometers themselves (comparing all others to one), the comparison results in an average error of 0.6 $\mu\mathrm{g\,m^{-3}}$ at 1 s resolution and 0.2 $\mu\mathrm{g\,m^{-3}}$ at 1 min resolution (15 % and 5 % error, respectively).

The spatio-temporal nature of mobile data requires several data cleansing and processing operations involving the interpolation and correction of GPS data. The GPS signal is sometimes lost or inaccurate during the monitoring. Because the GPS data were mostly missing for only short periods (2.1 s on average), a linear interpolation was used to insert these data points. When visualized on a map, GPS data were often slightly off track (due to reflectance by large buildings). Because the exact route was known, GPS data were not filtered on quality (e.g. using the number of satellites), but the GPS data were projected on the streets based on shortest distance. Only the streets that are included in the routes were used for this projection. Contextual information about travel direction and street orientation were further used to select the correct street to project the data upon. In Chapter 7, where the exact route is not known, a more elaborate filtering and processing of the GPS data is performed.

The accuracy of the GPS location of the measurements is related to the quality of the GPS signal (GPS accuracy and projection accuracy) and the synchronization of the measurement device. The uncertainty of the GPS location is difficult to assess, as we have no reference to compare with. The GPS device (GlobalSat BU-353) specifies an accuracy of the horizontal position of 10 m (horizontal root mean square (RMS) error). At certain locations in urban environments, this can be larger due to reflection of signals by high buildings (observed from plotting the geographical data), which can lead to artefacts in the data (e.g. a spatial shift in measured concentrations). The accuracy of the perpendicular projection on the streets could not be quantified exactly. The overall accuracy is probably around 10 m. The synchronization of the micro-aethalometer was checked and adjusted to the GPS time every day and the deviation was never more than 1 s. Given the average driving speed of 3.2 $\mathrm{m\,s^{-1}}$, this will not have an important impact on the spatial accuracy.

5.2.4 Data analyses

A multi-scale approach was applied in this study to investigate spatio-temporal UFP and BC patterns in an urban environment, refining the temporal and spatial resolution in different steps. In fact this is done through different levels of data aggregation starting from the original data at one second resolution. The spatial entities used in this study were the routes, the individual streets and finally points within a street. For the data analysis at the street level, measurements were aggregated at street level based on the validated GPS data. A total of 23 streets and squares were included in this study, a selection is given in Table 5.1. The highest spatial resolution in this study was obtained by data aggregation to fixed points within a street (see for example, Figure 5.8). A spatial database containing fixed points each 10 m along both monitoring routes was used for this aggregation of UFP and BC measurements. The data aggregation was based on a Gaussian weighing function with a standard deviation of 10 m. To each of the fixed points a value was attributed based on a weighted average based on distance. A temporal aggregation at different time scales was also applied. For the daily time scale, data were aggregated per monitoring day. These data are not to be confused with daily averaged data because the period of data collection was a six hours period between 7 a.m. and 1 p.m. Data aggregation to the hour of the day was also used, and finally some analyses were done using the maximal temporal resolution of one second.

The statistical comparison of pollutant concentrations is based on a non-parametric Kruskal-Wallis test at a critical value (α) of 0.05 (Kruskal and Wallis, 1952). This

is a non-parametric alternative to a one-way ANOVA. The Kruskal-Wallis test investigates whether the sampled populations have the same median. In case of multiple comparison the Kruskal-Wallis test was Bonferroni adjusted to compensate for multiple comparison. Correlation analyses were made by auto- and cross-correlation functions.

Exposure of a cyclist is assessed by calculating the average exposure level for each run. The variation observed in the exposure values is attributed to different sources of variation using Analysis of Variance (ANOVA). A spatial analysis of peak exposure is made using the fixed geo-point database that was used for the spatial aggregation. Peaks were defined as concentrations higher than the 95th-percentile concentration of BC and UFP. In a circular area with a radius of 10 m around each point, the number of peak events was counted and normalized for the number of runs. A counting scale between 0 and >1 was used, where low values (< 0.1) indicate the absence of peak exposure at a given location, and high values (≥ 1) indicate a peak exposure at nearly every passage. All statistical analyses were done in R (R Core Team, 2013).

5.3 Results and discussion

5.3.1 Pollutant variability between days

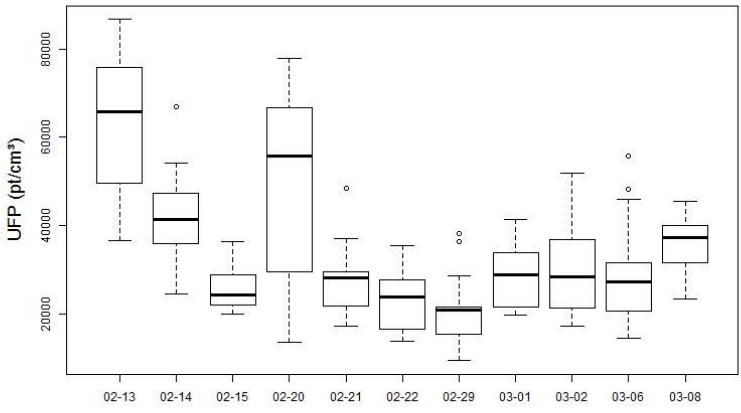
The mobile measurements were averaged per run for both routes. These averages were then compared between the different days of the monitoring campaign using Kruskal-Wallis tests. The UFP concentration showed a significant variability between the days (Chi-sq. = 159, d.f. = 10, $P < 0.05$, Figure 5.3a). The highest UFP concentrations were measured on February 13 and 20. During these two days, lowest average temperatures (2-4 °C, Table 5.2) in combination with highest relative humidity (89-95 %, Table 5.2) were observed.

The effect of meteorology on the daily averaged UFP concentration was significant for temperature (r = -0.74, $P < 0.01$), but not for relative humidity ($P = 0.50$) and wind speed ($P = 0.19$). This analysis is based on a limited dataset (11 daily averages) and therefore these conclusions are only valid for the study period and not necessarily for longer periods. However, studies over longer periods of time show temperature effects on UFP concentration between seasons (Padró-Martínez et al., 2012). Higher concentrations during winter months were attributed to higher PNC in vehicle exhaust and a greater atmospheric stability. Weichenthal et al. (2008) reported a decrease in PNC by approximately 10,000 pt cm^{-3} for each

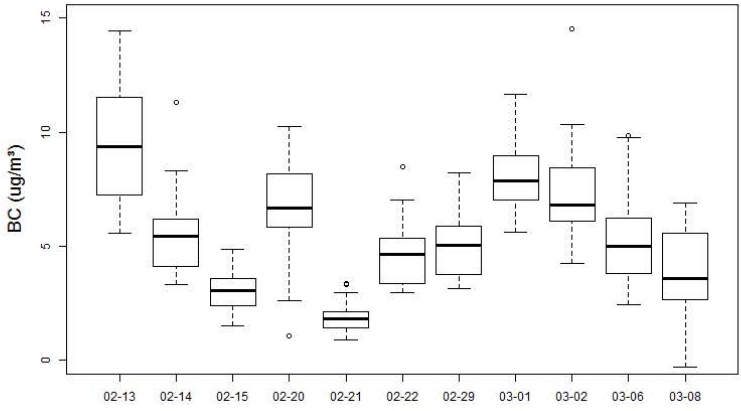
10 °C increase in morning temperatures. Low temperatures favour the formation of ultrafine particles in the atmosphere by cooling and dilution of semi-volatile gases from vehicle exhausts (Charron and Harrison, 2003). This effect may have caused the increased UFP concentration on the colder monitoring days of the study period.

BC concentrations along the monitoring routes did also change significantly between days (Chi-sq. = 213, d.f. = 10, $P < 0.05$, Figure 5.3b). A regression analysis between the averaged mobile BC concentrations and daily averaged data from three AQ monitoring stations at a traffic site (Borgerhout street side, BETR802, www.irceline.be), at an urban background site (Borgerhout backyard, BETR801) and at a rural background site (Dessel, BETN016) approximately 60 km far from the study area indicated a decreasing correlation (Figure 5.4). The slope of the regression lines decreased significantly for background data of Dessel, indicating that the contribution of rural background concentrations is relatively stable for the range of BC concentrations that were measured in urban environment during the study period. The urban background concentration and the measurements from the AQ monitoring station at the street side, in contrast, do account for a significant part of the variability in BC concentration obtained from mobile monitoring. From the AQ monitoring data, differences between street side, urban background and rural background concentrations are usually quite high. For some days, especially days when the BC concentration at the street side is low, the differences between the stations are much smaller. On these days the wind came from NE-E direction so that the traffic at Plantin en Moretuslei is downwind from the AQ stations. The contribution of the local traffic sources is lower on these days.

The correlation between BC concentration and meteorological variables is not significant for temperature ($P = 0.90$). Relative humidity ($r = 0.70$, $P < 0.05$) and wind speed ($r = -0.75$, $P < 0.01$), however, show a significant relationship with the BC concentration. These observations are valid for the study period but cannot be extrapolated due to the limited sample size and limited range of the meteorological variables. The inverse relationship between BC concentrations and wind speed indicates the importance of local primary emission sources. Low wind speed conditions cause less dispersion and as such a larger contribution of local emissions to the concentration levels. This is in line with results from literature studies based on larger datasets (Cao et al., 2009; Lewandowska et al., 2010; Maenhout and Cafmeyer, 1998). From a year-round monitoring study in urban environment, a strong inverse relationship between daily averaged BC concentrations and wind speed was observed by Viidanoja et al. (2002).



(a)



(b)

Figure 5.3: Boxplot analysis of the UFP (a) and BC (b) concentration for the 11 different measurement days.

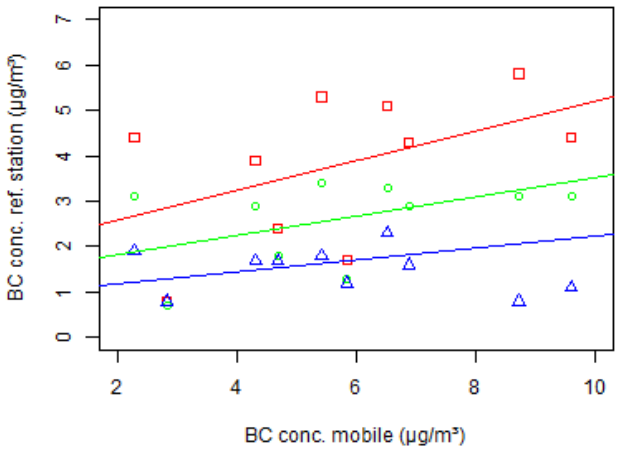


Figure 5.4: Regression analysis of the daily averaged mobile BC measurements and the daily averaged BC measurements at street side (red squares), urban background (green circles) and rural background (blue triangles) by stationary MAAP instruments.

5.3.2 Streetwise analysis of pollutant concentrations

The mobile data were aggregated at the street level by averaging measurements per street and per passage. As such, for each individual run, an average UFP and BC concentration was calculated for each of the streets. Previous case studies showed the potential of mobile monitoring for the assessment of spatial variability at street level and multi-street level (Peters et al., 2013b; Van Poppel et al., 2013). Over the entire monitoring campaign, significant differences were found in UFP concentration between the streets (Chi-sq. = 454, d.f. = 10, $P < 0.01$). The urban green Stadspark had the lowest UFP concentration, followed by low-traffic square (Dageraadplaats) and streets (Kleinebeerstraat, Table 5.3). The variability between the runs on the same day was limited in the urban green area, indicating that UFP concentrations did not fluctuate a lot over time on a given day. The UFP concentration in streets and squares of very low traffic density (e.g. Dageraadplaats and Kleinebeerstraat) showed a higher variability between the runs of the same day which is probably caused by some occasional vehicle passages. Carnotstraat, Plantin en Moretuslei, Quellinstraat, Bleekhofstraat and Provinciestraat had a high UFP concentration compared to the other streets. These streets are characterized by high traffic intensities (e.g. Plantin en Moretuslei, Table 5.1) and

a street-canyon topology (e.g. Provinciestraat). UFP concentrations at Wolfstraat showed a high variability between runs. Here the average UFP concentration was relatively high as could be expected based on the frequent bus traffic in this street, but high fluctuations were observed. The UFP concentrations were for some passages not significantly higher than at the urban green. In contrast, for other runs concentrations were similarly high as the concentrations in Carnotstraat or Provinciestraat and amongst the highest of the study period. This variation is explained by bus passages during the measurement runs causing high peak concentrations. At F. Rooseveltplaats which is a large square with a cross-road and bus stops, the UFP concentration was significantly lower than the UFP concentration in the street canyons, and significantly higher than in the urban green. This shows the effect of dispersion characteristics on concentration levels.

The BC concentration was also different between streets (Chi-sq. = 475, d.f. = 10, $P < 0.01$). The BC concentration at Dageraadplaats was significantly lower than at other streets. Stadspark had comparably low BC concentration. The highest BC concentrations were found in street canyons and traffic intensive streets. Large open areas such as F. Rooseveltplaats with busy cross-roads and bus stops had moderate BC concentrations, that were significantly lower than at street canyons with much less traffic (e.g. Provinciestraat).

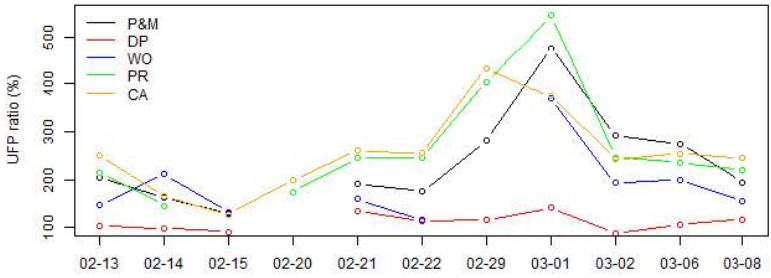
Table 5.3: Summary statistics of UFP and BC concentration in a selection of streets.

	UFP ($\mathrm{pt\,cm^{-3}}$)					BC ($\mu\mathrm{g\,m^{-3}}$)				
	Mean	Median	5th perc.	95th perc.	Sig.	Mean	Median	5th perc.	95th perc.	Sig.
Quellinstraat	45,757	42,962	22,008	83,314	a	8.3	7.0	2.8	19.1	ab
Provinciestraat	44,192	41,934	23,229	71,352	a	9.5	8.3	3.0	20.8	a
Carnotstraat	45,678	40,966	22,479	84,586	a	6.5	6.0	1.8	13.5	bc
Plantin en Moretuslei	43,07	38,525	19,675	82,45	ab	8.0	7.0	2.3	15.2	ab
Bleekhofstraat	39,874	33,479	18,36	77,272	ab	9.7	7.8	1.9	25.4	ab
F. Rooseveltplaats	35,198	32,3	15,085	55,448	bc	5.3	4.6	1.8	11.2	cd
Wolfstraat	36,463	26,541	11,383	85,431	c	8.1	5.5	1.3	24.5	bc
Lange Altaarstraat	33,977	27,297	11,972	75,481	c	6.1	4.3	1.1	12.9	d
Kleinebeerstraat	26,276	22,138	8,877	59,307	d	4.5	3.1	0.8	12.5	e
Dageraadplaats	22,4	19,787	8,754	48,03	de	3.7	2.7	0.9	6.3	f
Stadspark	19,793	17,493	7,924	39,543	e	3.2	2.9	1.2	6.2	ef
Overall	32,31	27,34	9,777	68,972		5.9	4.1	1.1	13.9	

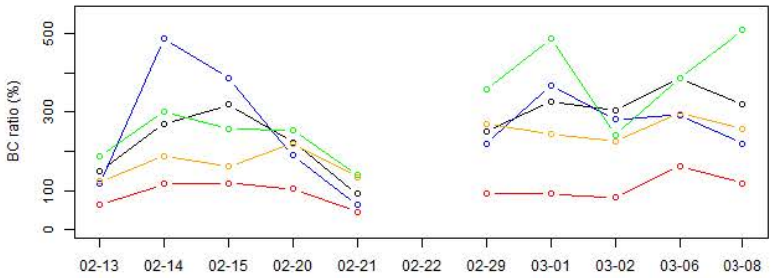
* Different letters indicate significant differences at the 0.05 level.

The ratio (expressed as %) between the UFP and BC concentration at a selection of streets and at the urban green (Stadspark) is given in Figure 5.5. Concentrations at the urban green are assumed to be at the urban background level because of the absence of local sources in the vicinity. Pollutant concentrations at a low-traffic square (Dageraadplaats) were comparable to the concentrations at the urban green for most days. For the other streets, considerable differences exist, and differences are clearly not stationary in time. The ratio for the street canyon of Provinciestraat, for example, ranges between 150 to 500 % on different dates for both UFP and BC, indicating the concentrations are 1.5 to 5 time higher than the urban background concentration. Obviously, this is related to significant local traffic sources in these streets. The fluctuations of the ratio for the different days may be caused by day-to-day differences in atmospheric conditions and dispersion and dilution processes of the local source emissions in addition to the variations in background concentrations (Figure 5.5c). The difference in exhaust emissions between the days is expected to be rather small for most of the streets due to the fixed measurement hours and inclusion of working days only. The difference in pattern of the UFP and BC ratios results from the different dynamics of both pollutants both in the background location and in the streets, i.e. the complex dynamics of condensation, coagulation and growth of UFP which is influenced by both temperature and atmospheric stability. Low ratios between street and background concentrations for BC and UFP were observed for 13/02, a day with very high background concentrations due to low temperatures and wind speed, i.e. high atmospheric stability and limited dispersion.

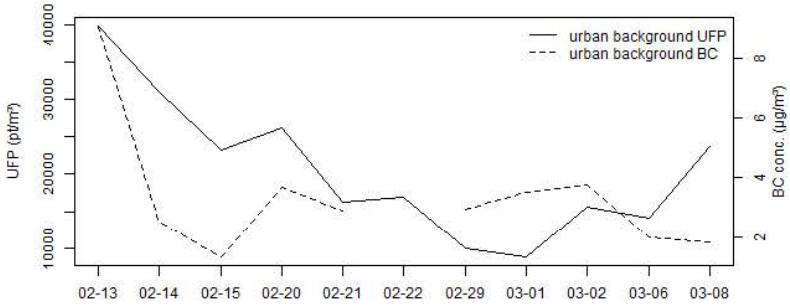
The significant differences in UFP and BC concentrations between the streets are partly related to the traffic density of the street. Traffic count data (Figure 5.6) show moderate correlation to averaged UFP and BC concentrations ($r = 0.57$ and $r = 0.23$, respectively). Franklin Rooseveltplaats, Plantin en Moretuslei and Carnotstraat show lower pollutant concentrations than what would be expected from the traffic counts, especially for BC. The main factors explaining this deviation are the distance to the traffic and the street topology. These three streets have their bicycle lanes separated from the traffic lanes, increasing the distance between cyclist and vehicles considerably. Additionally, F. Rooseveltplaats and Plantin en Moretuslei are more spacious with potentially increased dilution of pollutants. A stationary monitoring campaign performed in 2010 in the same urban area, showed that traffic load and also dispersion characteristics (distance and street configuration) affected spatial variability of UFP concentrations (Mishra et al., 2012).



(a)



(b)



(c)

Figure 5.5: Ratio between the daily averaged UFP (a) and BC (b) concentration at a selection of streets and at the background area (Stadspark). Daily averaged concentrations at the background location are given in (c).

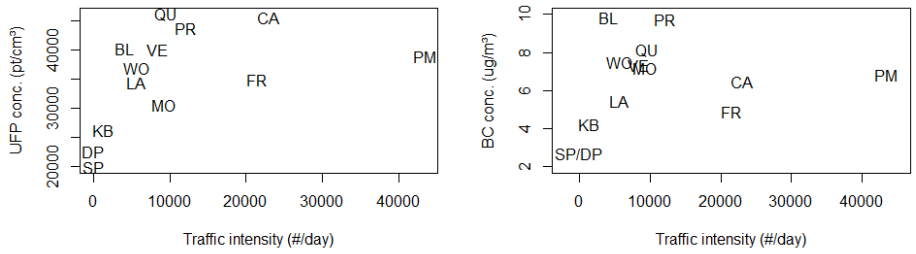


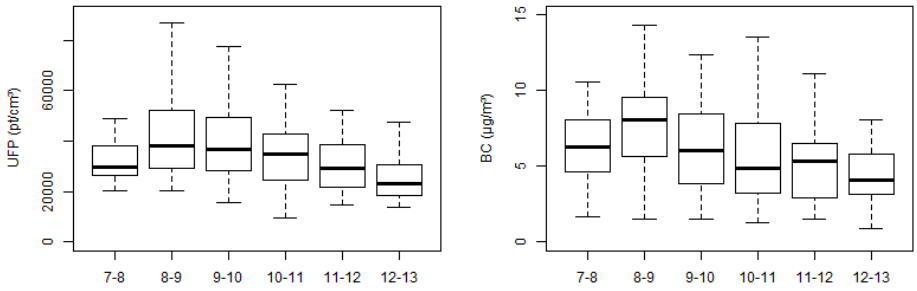
Figure 5.6: Average UFP (a) and BC (b) concentration per street in function of the traffic density. The labels are the abbreviated street names, abbreviations are given in Table 5.1.

5.3.3 Pollution variability within the day

Mobile data were averaged per hour of the day to obtain a diurnal pattern for the time frame of 7 a.m. until 1 p.m. (Figure 5.7). The UFP and BC concentrations showed significant differences between the hours of the day (Chi-sq. = 49, d.f. = 6, $P < 0.01$, and Chi-sq. = 38, d.f. = 6, $P < 0.01$). The highest concentrations were measured between 8 and 9 a.m., corresponding to the morning rush hour. As an example, vehicle counts at Plantin en Moretuslei increase from 1018 between 6–7 a.m., to 2020 between 7–8 a.m. and 2531 between 8–9 a.m., and decrease to 2200 vehicles over the following hours (validated count data from 2010, not published).

5.3.4 High-resolution patterns

A spatial aggregation algorithm (Gaussian distance weighting function) was used to calculate UFP and BC concentrations at a 10 m spatial resolution along both routes. The resulting maps based on all the measurements from the study period showed considerable variability (Figures 5.8 and 5.9). Main observations are similar to the streetwise analysis presented: low concentrations at background areas and high concentrations at high-traffic streets and at locations close to sources and with low dispersion characteristics. However, some additional observations can be made. The traffic-free urban green area (Stadspark) has a significantly lower UFP concentration, even close to the border of the park. This suggests a rapid decline in UFP concentration with distance from the traffic sources. Other



(a) (b)

Figure 5.7: UFP (a) and BC (b) concentration pattern over the period 7 am until 1 pm.

streets with low UFP concentration were the traffic-free square Dageraadplaats, a dead-end street (Korte Altaarstraat) and other small streets in the same residential area. The spatial variability of UFP in these streets is also low. The road with the highest traffic density of the study area (Plantin en Moretuslei) showed two distinct parts (PM-A and PM-B, see Figure 5.8). One part (PM-A) has a UFP concentration in the range of 30,000 – 50,000 pt cm^{-3} with significant variability along the road. The second part, more towards the city centre (PM-B), has a UFP concentration in the range 50,000 – 80,000 pt cm^{-3}. The highest values are observed in a short tunnel (70 m) underneath the railways. The major difference between both parts of Plantin en Moretuslei is the distance of the biking lane from the road, with a considerably larger separation between biking and driving lane for part PM-A (5 m, compared to < 2 m for part PM-B). The street canyons without separate bicycle lane (part of Provinciestraat and Quellinstraat) have high UFP concentrations along the entire street length.

Similar trends are observed for BC. The highest concentrations are obtained in the tunnel of Plantin en Moretuslei (up to 17 $\mu g\, m^{-3}$), the lowest in the urban green (3 $\mu g\, m^{-3}$). The BC concentration decreased rapidly from 6 $\mu g\, m^{-3}$ at the entrance of the park near the street to 3 $\mu g\, m^{-3}$ at a distance of about 25 m from the street. The dilution effect is likely to cause the difference between the two parts of Plantin en Moretuslei as well. The streets with low traffic density in the neighbourhood of Dageraadplaats have a much lower BC concentration than other neighbouring streets with a higher traffic density. This 10 m resolution map shows potential for hot-spot location identification to support traffic measures

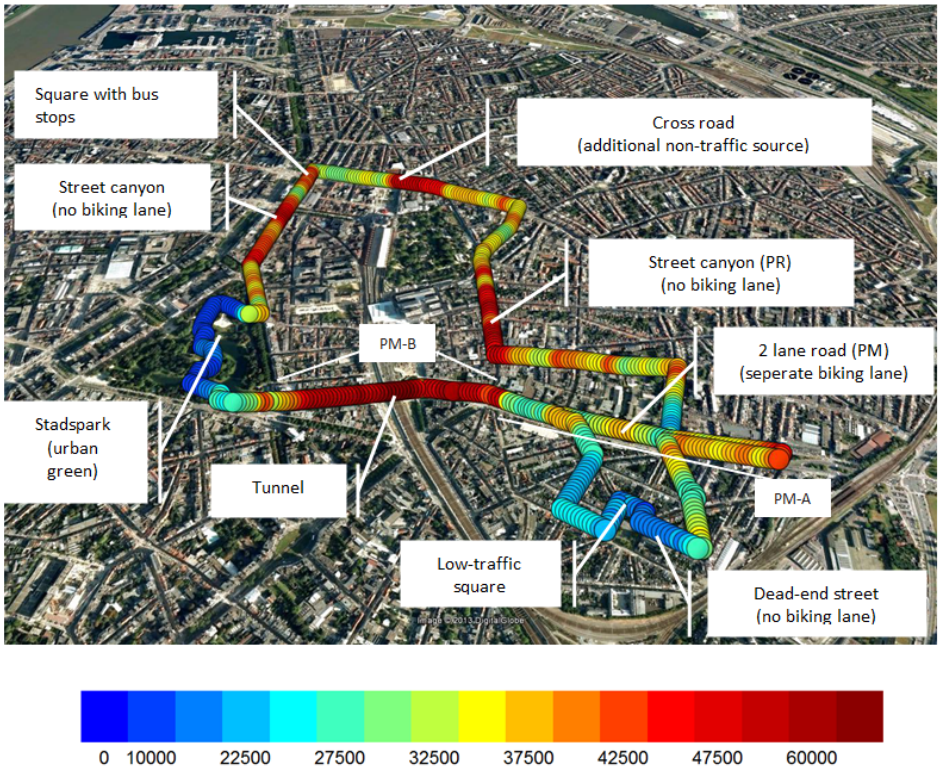


Figure 5.8: Map of the UFP measurements aggregated to a 10 m spatial resolution.

(e.g. location of biking lanes in relation to traffic numbers as function of street characteristics).

Spatial pollutant patterns could be distinguished for different hours of the day (see Suppl. 5.S2-5.S3). At the start of the morning rush (7–8 a.m.), UFP and BC concentrations are moderate in general, but some localized hotspots are already distinguished (for example, near the tunnel on Plantin en Moretuslei). The UFP concentration at the urban green is low on this hour of the day ($< 15{,}000$ pt cm^{-3}). Between 8 and 9 a.m., the UFP concentration increases significantly, and hot-spots expand spatially, especially in street canyons and between intersections at busy roads. After 9 a.m., an opposite decreasing trend in UFP concentration was observed and the spatial extent of hot-spots decreases in the hours between 9 a.m. and 1 p.m. In some street canyons (e.g. Quellinstraat, part of Provinciestraat), however, UFP concentrations stay high, probably due to a reduced air circulation in these streets or street-specific source dynamics.

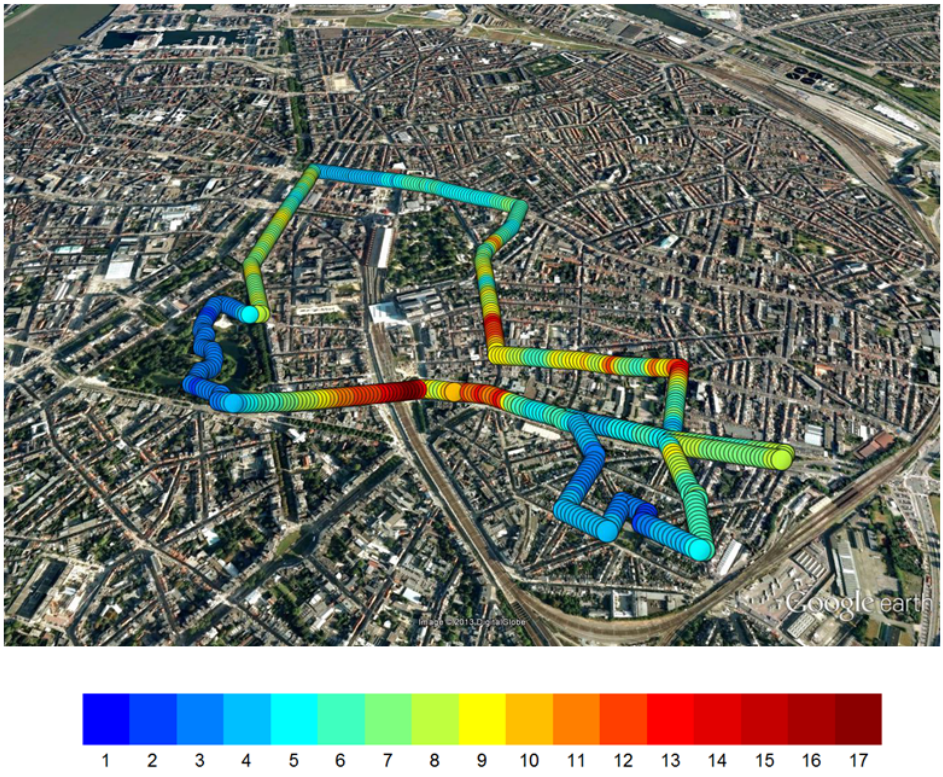


Figure 5.9: Map of the BC measurements aggregated to a 10 m resolution.

Table 5.4: Averages (standard deviation) of the correlation between UFP and BC for each individual run for different streets.

Route 1		Route 2	
Street name	Correlation	Street name	Correlation
Plantin en Moretuslei	0.61 (0.31)	Plantin en Moretuslei	0.48 (0.32)
Kleinebeerstraat	0.45 (0.24)	Stadspark	0.34 (0.51)
Dageraadplaats	0.30 (0.29)	F. Rooseveltplaats	0.32 (0.35)
Korte Altaarstraat	0.43 (0.22)	Carnotstraat	0.25 (0.30)
Wolfstraat	0.59 (0.18)	Provinciestraat	0.49 (0.39)

5.3.5 Correlation between UFP and BC

The Pearson correlation between the spatio-temporally aggregated BC and UFP data at a 10 m resolution over the entire monitoring campaign was high (r of 0.82, $P < 0.01$), indicating that the spatial concentration pattern that cyclists are exposed to in urban environments is very similar for UFP and BC. Despite their different dynamics related to meteorological conditions (Section 5.3.1), UFP and BC are both strongly traffic related pollutants. This explains that the hot-spots with high UFP and BC concentrations coincide largely. Further, the correlation between UFP and BC was calculated for the individual runs. Averages and standard deviations of these correlations are given in Table 5.4 for a selection of streets. The correlation between UFP and BC was not significantly different between all the streets at both routes ($P = 0.38$ and $P = 0.41$, for route 1 and 2, respectively). So, at traffic-free and calm streets the correlation between UFP and BC is not significantly different compared to streets where UFP and BC levels are elevated. The lower correlation at Carnotstraat is caused by a systematic UFP peak from a non-traffic source (cooking and frying activities at a restaurant) at the entrance of the street.

5.3.6 Cyclist exposure

Mobile monitoring allows to collect air quality data at a high spatial resolution. As seen, these data can capture local scale variations in exposure which makes them potentially suitable for detailed exposure assessment. The average exposure of a cyclist was calculated for each of the runs of route 1 and 2 separately and integrates streets of different topology, traffic density and pollution levels. The number of runs was reduced based on the time needed to complete a given run; 206 runs were included in the analysis for route 1 and 82 for route 2. All the

Table 5.5: Summary statistics of the average exposure of a cyclist to UFP (in $\mathrm{pt\,cm^{-3}}$) and BC (in $\mu\mathrm{g\,m^{-3}}$) during a short trip in urban environment.

	Route 1		Route 2	
	UFP	BC	UFP	BC
Min.	3,800	1.3	16,000	1.9
1st. Qu.	17,300	3.3	28,500	4.6
Median	23,000	4.9	36,500	7.2
Mean	27,100	5.7	37,600	7.3
3rd Qu.	31,900	7.7	42,100	9.1
Max.	87,300	14.1	77,900	18.0

selected runs of route 1 were driven in 8 to 13 minutes, those of route 2 in 21 to 32 minutes. The runs that were deleted from this analysis included a large number of non-mobile measurements when the cyclist stopped for long periods at specific locations.

Exposure values showed a high variability between the runs (Table 5.5) for both pollutants. Values were higher for route 2. The average exposure of cyclists to UFP and BC is influenced by the day-to-day variation of pollutant concentrations, the UFP and BC concentration pattern within the day and the occurrence of peak events along the cycling route. For short trips, all of these factors are significant in explaining the variations in cyclist exposure (ANOVA, $P < 0.01$). Nevertheless, the occurrence of peak events (UFP $> 80000\ \mathrm{pt\,cm^{-3}}$, BC $> 15\ \mu\mathrm{g\,m^{-3}}$) is the dominant factor. The exposure of cyclists to local peaks of UFP and BC in an urban environment occurs for example at crossroads, by passing (heavy duty) vehicles, or when the cyclist is driving behind a vehicle for a period of time. The former cause is more structural, whereas the latter two are more coincidental. A spatial analysis of the locations of peak exposure is presented to assess the importance of structural peak exposure hot-spots (Figure 5.10), where at locations with high values (≥ 1) the cyclist is exposed to peak concentrations in almost every passage. Peak exposure mainly occurs at (major) cross-roads (Figure 5.11 A-C,F) and in a tunnel (Figure 5.11 C). Peak exposure is absent in urban green areas (Stadspark, Figure 5.11 D) or at a low traffic square (Dageraadplaats), but also in Carnotstraat (Figure 5.11 E), a street canyon with high average concentrations of UFP and BC (see Table 5.3). The street topology with a separate biking lane and the relatively smooth traffic flow without stops can explain the absence of peak exposure events at Carnotstraat. This does of course not mean that exposure in Carnotstraat is low (concentrations are high), but only that the contribution to the

overall exposure in Carnotstraat due to peak events is limited, in contrast to the high contribution of peak events when the whole biking route is considered.

Given the localized nature of peak events – both in space and time – and its primary contribution to the exposure to pollutants for cyclists, exposure estimations are preferably based on high-resolution data. Many examples in the literature illustrate the use of personal monitoring, where selected individuals from a studied population are equipped with monitoring devices to gain insight in their exposure to air pollution (e.g. Dons et al., 2012). For these individuals, the actual exposure to a given pollutant is measured concurrently with data on position, transport mode or other parameters, and provide precise information on the exposure history of (a small) number of individuals. Recently, methods have been developed to use ubiquitous technology for exposure assessment of larger populations. de Nazelle et al. (2013) used smartphone technology to obtain geospatial, time and activity data, and combined this information with air quality maps to assess the exposure to air pollution. The authors highlight the potential of this technology to unobtrusively monitor the exposure of large groups of individuals at low cost. The accuracy of this approach depends greatly on the accuracy of the air pollution maps where the trajectories are projected upon. Mobile monitoring may potentially be used to generate high-resolution cyclist exposure maps over larger (urban) areas. However, several issues can interfere with the real-life applicability of the type of mobile monitoring that is presented in this study. Firstly, to generate high-resolution cyclist exposure maps over larger areas, repeated monitoring runs are needed to map localized patterns at sufficient detail through time (Peters et al., 2013b), which is very time consuming in the current set-up. This will be further investigated in the next chapter. Furthermore, these cyclist exposure maps might not be representative for other transport modes. Several studies (e.g. de Nazelle et al., 2012; Dons et al., 2013a) have demonstrated the impact of transport mode on exposure, but the potential of extrapolating mobile measurements from one transport modus to another has, as far as we know, not been studied in sufficient detail. In other words, there is still a need to actually measure in different transport modes, further increasing the mobile monitoring efforts.

5.4 Conclusion

The urban environment shows a high variability in UFP and BC concentrations. The density and dynamics of the traffic and other sources in conjunction with the location-specific mixing and removal mechanisms results in a complex pollution pattern of urban micro-environments. The spatial variability of UFP and BC

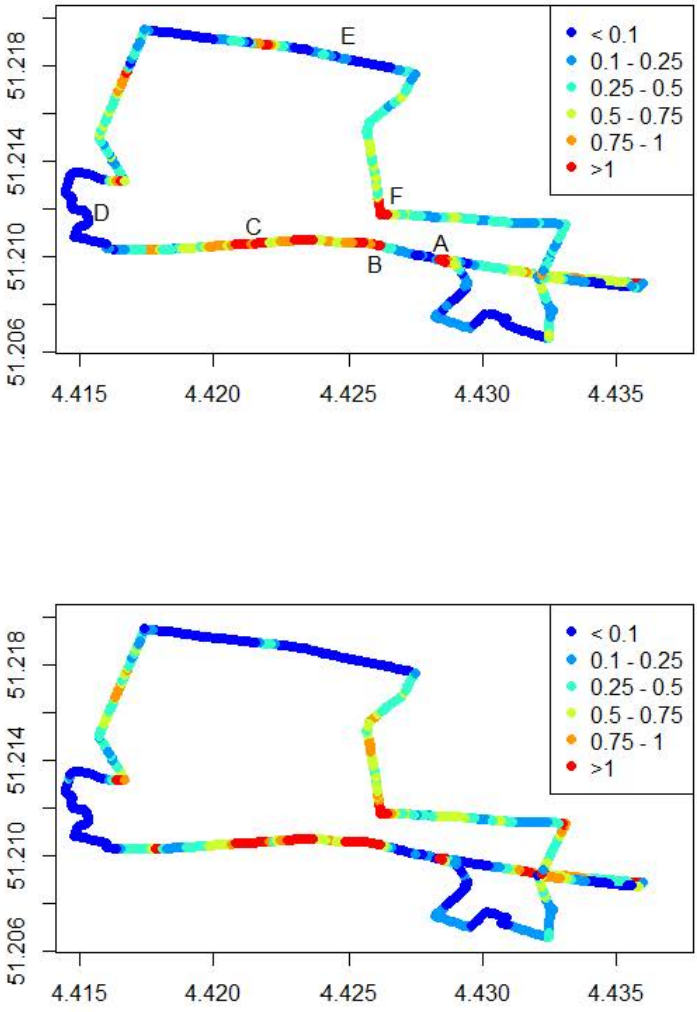


Figure 5.10: Occurrence of peak exposure along the bicycle routes for UFP and BC expressed as the number of peak UFP and BC measurements normalized for number of repeated runs.

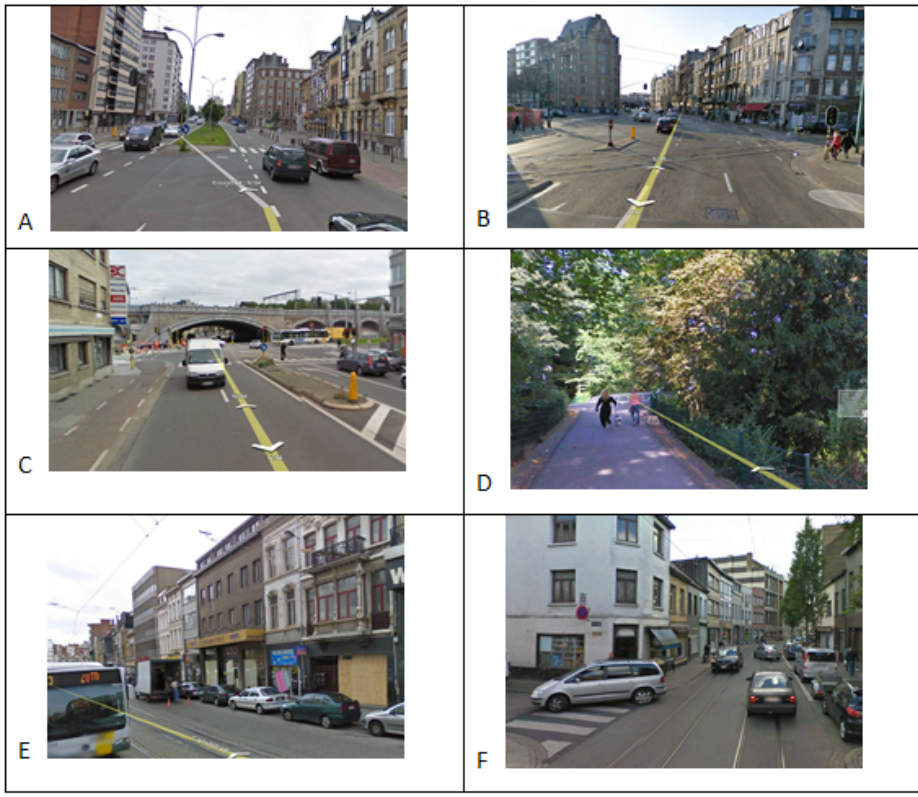
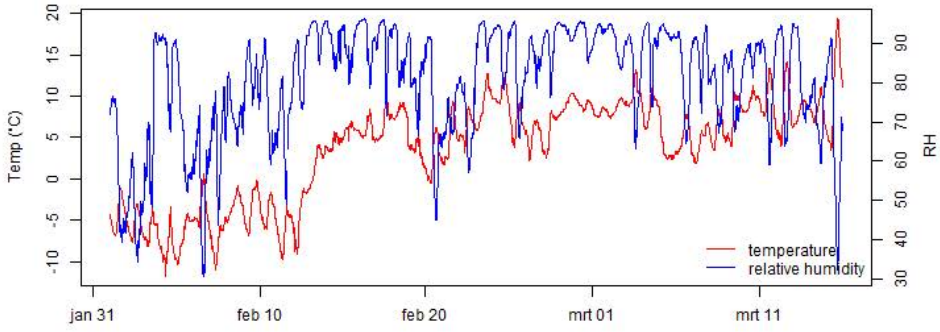


Figure 5.11: Illustration of the locations indicated in Figure 5.10.

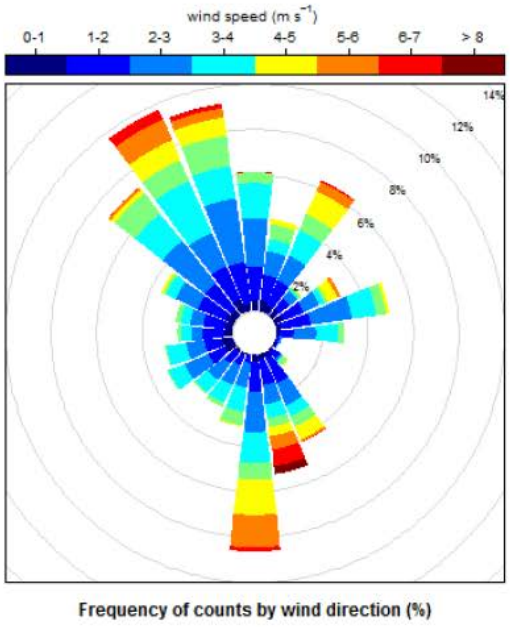
concentrations showed a relation with traffic intensity at the urban study sites, however this was not the only explaining variable of pollution levels: also street topology and distance to traffic had a significant impact. Following the variations in pollution concentrations, integral exposure of cyclists to UFP and BC is highly variable in time in urban environment as well. At the urban study site, the primary factor determining the integral exposure during the bicycle trips was the occurrence of local peak concentrations, both for UFP and BC. Variations between days and within the day were also significant, but less important. These results underline the significance of local contributions to the air pollution, and suggest that cycling infrastructure and management has the potential to reduce the exposure to air pollution for cyclists by avoiding busy streets with high pollution levels and hot-spots with peak concentrations.

A confined and targeted monitoring campaign was used to perform spatio-temporally resolved comparisons of exposure of cyclists to UFP and BC concentrations at various levels of detail. This chapter showed that mobile monitoring can be used to map the variability in urban air quality at a high spatial resolution and that it can give additional insights in exposure assessment. In the next chapter, a more detailed look is taken at the methods for mapping the air quality using mobile monitoring using the dataset described in this chapter.

5.A Supplementary data



(a)

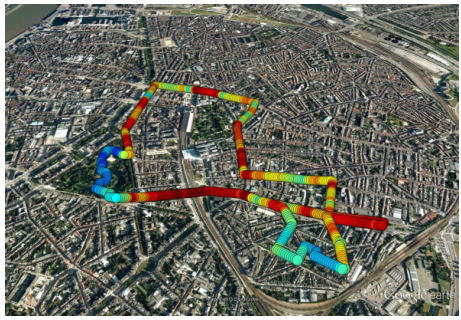


(b)

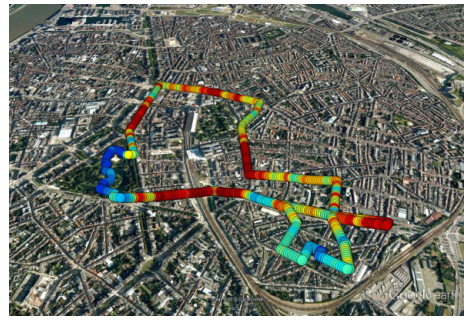
Figure 5.S1: Temperature, relative humidity (a) and wind conditions (measured at a height of 30 m) (b) in the winter of 2012 (2012, Feb 1 Mar, 15).

(a) 7 h–8 h

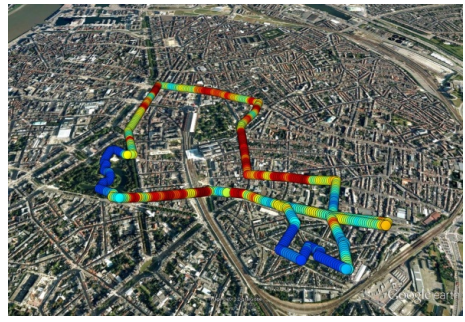
(b) 8 h–9 h



(c) 9 h–10 h



(d) 10 h–11 h



(e) 11 h–12 h

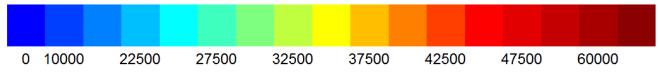
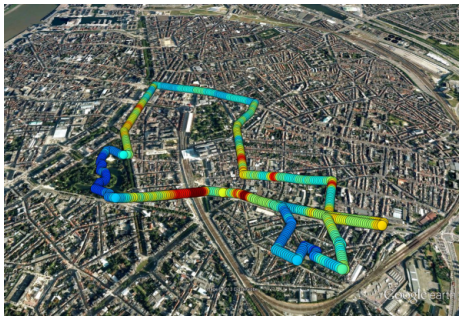


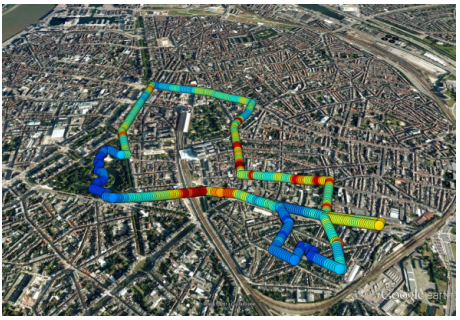
Figure 5.S2: UFP concentration at different hours of the day

(a) 7 h–8 h

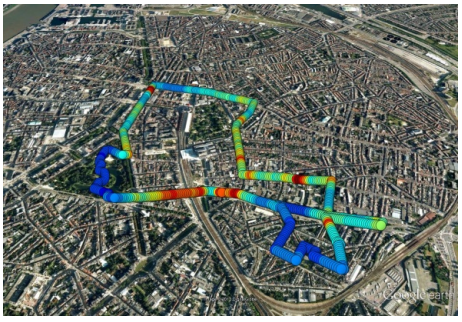
(b) 8 h–9 h



(c) 9 h–10 h



(d) 10 h–11 h



(e) 11 h–12 h

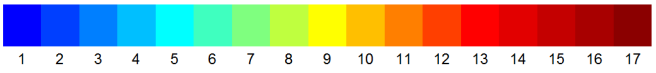


Figure 5.S3: BC concentration at different hours of the day

CHAPTER 6

Development and validation of a mobile monitoring methodology based on an extensive dataset

6.1 Introduction

In contrast to traditional stationary monitoring stations, mobile platforms are able to acquire air quality data at a high spatial resolution (Wallace et al., 2009; Zwack et al., 2011b). But at the same time, due to the high temporal variability of the urban air quality and the mobile nature of the measurements, the representativeness of the mobile measurements is a major issue. These considerations are very relevant given the increasing use of mobile air quality monitoring as a solution to measure micro-scale variability at a high spatial and temporal resolution. Kuhlbusch et al. (2014), for example, mention mobile monitoring to collect highly spatially and temporally resolved data in their recommendations for future European air quality monitoring.

The objective of this chapter is to develop and validate a method to map urban air quality at high spatial resolution using mobile monitoring. Based on a large experimental dataset of mobile air quality measurements (presented in the previous chapter), the impact of the high spatio-temporal variability is investigated,

This chapter is based on the following publication: Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J. and De Baets, B. (2015) Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. *Atmospheric Environment*, 105 148-161.

leading to practical considerations and proposed guidelines with regard to mobile monitoring campaigns to map the urban air quality.

Some critical issues related to mobile monitoring of air quality have to be acknowledged. These issues arise from the combination of the high temporal variability of air quality and the mobile nature of the measurements (see Section 4.3). When mobile monitoring is used to derive high resolution maps that are representative in time, large amounts of data are required to represent the range of possible meteorological and traffic conditions (Padró-Martínez et al., 2012) and data aggregation has to be performed. Peters et al. (2013b) indicate the need for repeated measurements to map local pollution patterns. In the previous chapter, we demonstrated that the variability in urban BC and ultrafine particles (UFP) concentrations can be mapped at a spatial resolution of 10 m. However, up to 256 repeated measurement runs were used, which is not always practically feasible. Van Poppel et al. (2013) applied background correction to reduce the number of repetitions required to obtain representative results of the spatial variability of pollutants at different micro-environments in a city. The results are also sensitive to the chosen data processing approaches, as stressed by Brantley et al. (2014b), who explored the influence of different techniques for local emission event detection, background estimation and averaging.

Therefore, suitable monitoring strategies and analysis methods have to be defined that take spatio-temporal representativeness into account. The large number of repeated measurement runs performed in this study allows to address the following research questions about obtaining a representative and consistent image of the spatial variability of the urban air quality: (i) What is the impact of events and the used spatial resolution? (ii) How representative is this image in time? (iii) How many runs are required to obtain this image of the spatial variation? (iv) Can we reduce the required number of runs by reducing the variability through filtering for events or by applying background normalisation?

6.2 Material and methods

6.2.1 Mobile monitoring campaign

An extensive monitoring campaign using a mobile monitoring platform, the Aeroflex, was set up in Antwerp, Belgium, during 4 weeks in February-March 2012, and described in detail in the previous chapter. For more details on the set-up of the monitoring campaign, see Section 5.2.1. In this chapter, we will only make use

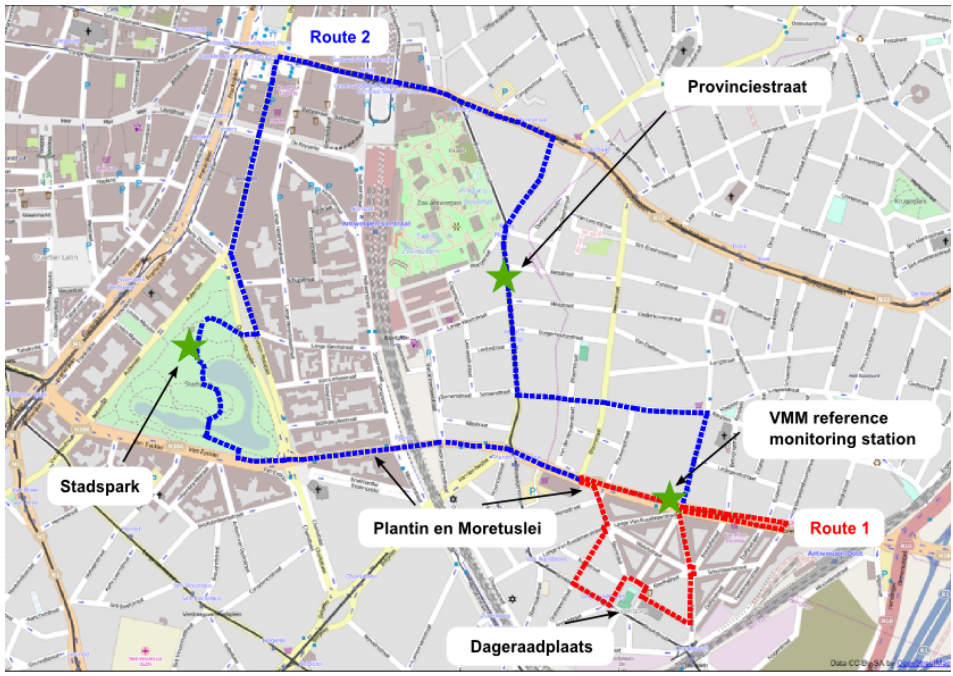


Figure 6.1: Overview of the bicycle routes in the city of Antwerp, Belgium: route 1 (red) and route 2 (blue). The green stars indicate the locations of the stationary measurements.

of the BC data. A description of the data validation and quality control of the BC and GPS data is given in Section 5.2.3.

A controlled monitoring set-up with fixed routes and time slots was used to guarantee sufficient temporal and spatial coverage to address the research questions. The mobile measurements were made along two fixed routes (about 2 and 5 km long, see Figure 6.1). An overview of the most important streets of both routes is given in Table 5.1 in the previous chapter.

The campaign consisted of 256 runs along route 1 and 96 runs along route 2, spread over 11 days in a 4-week period between 2012-02-13 and 2012-03-08. The runs always occurred between 7 am and 1 pm on weekdays, covering the morning rush, and were performed with two equipped bikes, monitoring the two routes simultaneously. This resulted in the collection of 92 h of mobile measurements.

6.2.2 Complementary stationary measurements

Stationary measurements at 3 locations accompanied the mobile campaign. BC was measured continuously during the 4 week period of the mobile campaign at the Stadspark (urban green), Provinciestraat (street canyon) and Plantin en Moretuslei (two times two lanes entrance road). See Figure 6.1 for the exact locations. These stationary BC measurements were performed with micro-aethalometers with a flow rate of 50 $\mathrm{mL\,min^{-1}}$ and a 5 min measurement rate.

Furthermore, measurements from the VMM reference monitoring station at the Plantin en Moretuslei were also available. At this station, BC is measured with Multi-Angle Absorption Photometry (MAAP, Petzold and Schonlinner, 2004; Petzold et al., 2005) at 1 min resolution.

6.2.3 Data analysis and processing methods

The term *run* is used to denote the completion of an entire route once, and this is repeated at different days and/or different times of one day. During each run, a dataset of point measurements is collected. When the mobile platform passes by at a certain location during a run, this is called a *passage*. In this way, the monitoring campaign consists of a number of passages at the different locations of a route.

The analyses in this study are performed at different spatial scales by spatially aggregating the original data at different levels. The spatial entities varied from the entire routes, the different individual streets of a route, to segments of a certain length within a street. The street segments were constructed at different spatial resolutions (segment length varying from 5, 10, 20, ... to 100 m) along the streets by dividing each street section (from one intersection to another) in segments of equal length of approximately the spatial resolution. The mobile data were spatially aggregated by allocating the measurements to one of the spatial entities based on the projected GPS data. The measurements allocated to a specific spatial entity are then aggregated by computing the arithmetic mean per passage, i.e. the consecutive 1 s measurements are combined per spatial entity. As such, for each individual run, an average BC concentration was calculated for each spatial entity, resulting in *passage means*.

For the highest spatial resolution, i.e. the segments (represented by their midpoints), concentration profiles were constructed to explore the spatial variability at this level. On these figures, the mean concentration and standard error on the mean, calculated for each segment based on the passage means, are plotted.

Different methods are explored to reduce the impact of extreme values on the spatially aggregated result. Firstly, two other averaging statistics are used besides the arithmetic mean: the median and the α-trimmed mean. The α-trimmed mean is computed by removing the α largest and α smallest values and computing the arithmetic mean of the remaining values. This is a statistical measure of central tendency that is less influenced by the effects of the tails of a distribution. As the air quality data in this study are positively skewed, removing the tail measurements will result in a biased estimator of the population mean. Indeed, this estimator will systematically underestimate the population mean. However, due to the reduced tail effects we believe this is an appropriate statistic. In this study, a trimmed mean resulting from the removal of the lowest and highest value is adopted. This corresponds to $\alpha = 0.5\,\%$ for samples of up to 200 runs. Unless specified otherwise, trimming was always performed on the passage means and for each spatial entity separately (the passage mean itself is not calculated with trimming). Secondly, an emission event filtering method is used: the running coefficient of variation (COV) method from Hagler et al. (2012). For this method, a running 5 s standard deviation of the BC concentrations is calculated and divided by the mean concentration of the entire sampling period. The 99th percentile of the calculated COV is used as a threshold and all data points with a COV above this threshold are removed along with the data points 2 s before and after (Brantley et al., 2014b).

Background concentrations are obtained from the stationary measurements at the centre of an urban green (the Stadspark) without sources within a direct vicinity of 50 m. These measurements can therefore be considered as urban background concentrations (Hoek et al., 2002). These measurements are available at 5 min resolution and are processed in two ways: the hourly median value and the moving mean (using a Gaussian window with standard deviation of 30 min). The mobile measurement data are normalised with a background normalisation method based on these values to minimize the influence of meteorological day-to-day variations in the urban air quality. This normalisation is performed through a combination of the additive and multiplicative method as used in Dons et al. (2012). The normalised values are calculated as:

$$\mathrm{BC}_{norm,i} = \mathrm{BC}_i - \mathrm{BC}_{bg,i} + \overline{\mathrm{BC}_{bg}} \qquad \text{(additive)} \tag{6.1a}$$

$$\mathrm{BC}_{norm,i} = \mathrm{BC}_i / \mathrm{BC}_{bg,i} \cdot \overline{\mathrm{BC}_{bg}} \qquad \text{(multiplicative)} \tag{6.1b}$$

with BC_i the original BC measurement at time i, $\mathrm{BC}_{bg,i}$ the background measurement at time i, and $\overline{\mathrm{BC}_{bg}}$ the mean background concentration for the full period. The additive method is applied to high concentrations (measurement is greater than the background value) and the multiplicative method to lower concentra-

tions (measurement is smaller than the background value), see Dons et al. (2012) for more details.

Analyses were performed using R (R Core Team, 2013), pandas (PyData Development Team, 2014) and QGIS (QGIS Development Team, 2014).

6.2.4 Data experiment

A data experiment was conducted to investigate the representativeness of the measurements, i.e. to address the research question of how many mobile runs are required to obtain a representative estimate of the air quality in a certain area. In this study, a representative estimate means that the average BC concentration of a *subsample* of runs for a spatial entity is within a certain percentage, the *deviation* (e.g. $\pm 25\,\%$), from the overall average value obtained from the entire mobile measurement campaign for that spatial entity. In the absence of an integrated average of continuous measurements at each location, this overall average is considered as a good approximation of the pollutant concentration to be measured. The representativeness is thus constrained by the space and time of the data collection, which was biased towards daytime hours during one month of the year in an urban environment, which is not necessarily representative for other periods or locations.

The data experiment consisted of repeatedly taking subsamples through the generation of random combinations of all runs. Starting from the measurements of a single run and then cumulatively adding measurements from randomly selected runs (random selection with replacement), the average concentration was calculated for each combination of increasing size for each spatial entity. This way, the average pollutant concentration could be evaluated in function of the increasing number of runs included in the sample and compared with the overall average value to see after how many sampled runs the average converged to the overall average. Convergence is obtained when the average of the sampled runs deviates less than the deviation percentage from the overall average, and does so consistently when adding new runs. For one experiment, the procedure explained above is repeated a high number of times (1000 iterations, where each iteration is one random combination of the runs) to minimize the influence of random effects. This data experiment is based on the work of Peters et al. (2013b), but differs from it in the application of a random selection with replacement instead of without replacement. When using random selection without replacement (each run can only be selected once), the average would converge exactly to the overall average when more runs are added and convergence would always be obtained. To prevent this,

random selection with replacement is used. The data experiment was repeated for different set-ups: at different levels of spatial aggregation, with the arithmetic and trimmed mean, and with and without background normalisation.

6.3 Results

6.3.1 Occurrence and impact of events

Mobile air quality measurements as described in this study show typical characteristics due to the high temporal resolution of the data collection and the varying spatial context. A large variability is observed in the measurements at one second resolution of one run (Figure 6.2a). Measurements with a low variability (e.g. Stadspark) are followed by measurements with a higher variability in other streets. In the regions with a high variability, sharp peaks are seen due to the proximity to sources, leading to large differences in BC levels within a couple of seconds. Events, typically a closely passing car or bus, cycling in the emission plume right behind a vehicle, idling vehicles or local traffic congestion, are causing these high BC levels. These events can occur systematically at certain places or can be more accidental. We will refer to these last events as random events. The occurrence of events leads to a distribution of the data as shown in Figure 6.2b. A part of the BC range is cut off on this figure: around 0.5 % of the data at one second resolution fall outside the range of the figure (0-50 $\mu g\,m^{-3}$). This highest 0.5 % of the measurement data has a moderate impact on the overall mean of all runs: discarding these values decreases the overall mean with 7 %. However, the following two examples will illustrate that extreme values can have a much larger impact at a smaller spatial resolution or when having fewer repetitions.

The first example is the distribution of the mean BC concentration per passage at the street level in the Grotebeerstraat (Figure 6.3). Some passages have a clearly higher mean than most others, with a maximal value of 33.1 $\mu g\,m^{-3}$. Such an extreme value can have a big impact on the resulting average street level concentration of the Grotebeerstraat. On all 256 runs, this influence is limited (a mean of 3.6 $\mu g\,m^{-3}$ for all runs vs. 3.5 $\mu g\,m^{-3}$ without the extreme). However, in the case of a much smaller sample size, e.g. 20 randomly selected runs, this impact is very significant: 5.2 vs 3.5 $\mu g\,m^{-3}$ with and without the extreme, which is an increase of around 50 % due to a single passage.

The second example is shown in Figure 6.4, which depicts the concentration profile in the Provinciestraat at 20 m resolution. The mean concentration profile is shown

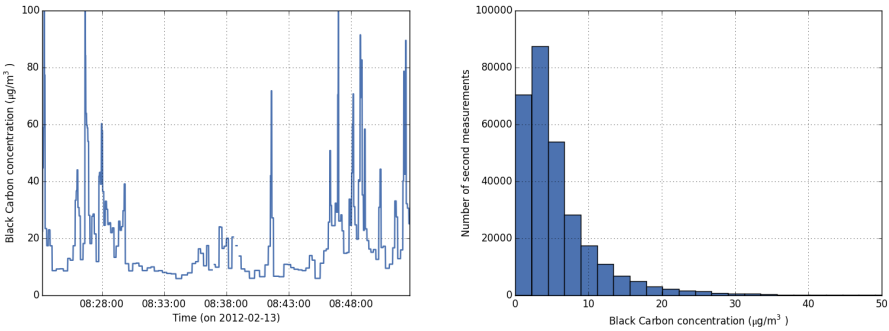


Figure 6.2: (a) Example run of mobile measurements at one second resolution for route 2 on 2012-02-13 and (b) the distribution of all mobile measurement data at one second resolution for routes 1 and 2 combined.

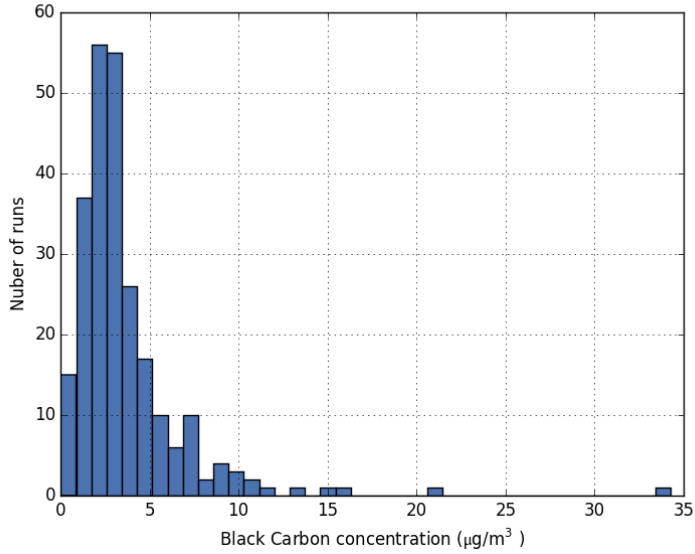


Figure 6.3: Distribution of the passage means for the Grotebeerstraat (route 1, a total of 249 runs included).

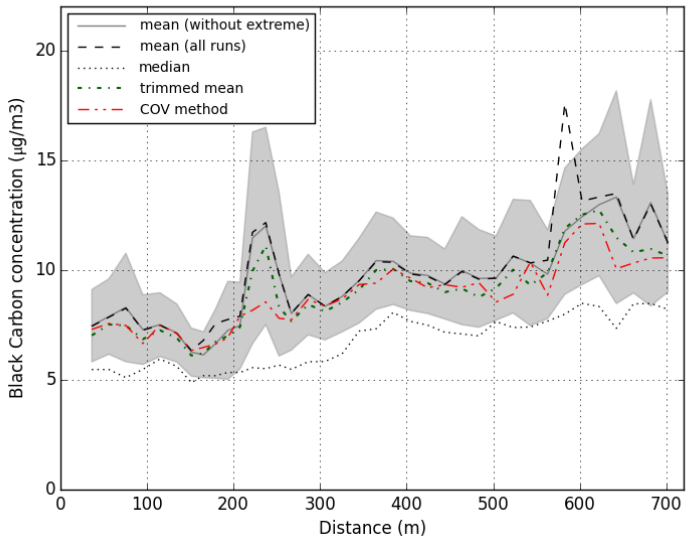


Figure 6.4: Concentration profile of the Provinciestraat: BC concentration levels at 20 m resolution using different processing methods: mean, median, trimmed mean and COV method. Shaded area depicts two times the standard error on the mean (for the mean without extreme).

for all 96 runs (dashed line), and there is a peak in the BC concentration at 225 and at 580 m. Looking at the peak at 580 m, it appears that this peak can largely be attributed to a single measurement run that includes extreme measurements. Excluding this particular run smooths the BC profile and lowers the average BC concentration at that place from 17.6 $\mu g\,m^{-3}$ to 11.8 $\mu g\,m^{-3}$ (a reduction of 49 %). If the total number of runs is 20 instead of 96, the impact is even larger: 43.3 vs. 12.1 $\mu g\,m^{-3}$ for 20 randomly selected runs with and without the extreme, which is an increase of more than 250 % due to a single passage. The previous examples show that a single run can have a large impact on the result, especially at a smaller spatial resolution or when having fewer repetitions.

Next, different aggregation functions besides the arithmetic mean are explored to decrease the impact of these extreme measurements, while retaining the characteristic concentration profile: median, trimmed mean and the rolling COV method (see Section 6.2.3). The results of these methods can be seen in Figure 6.4 for the Provinciestraat. With all the different statistics, the concentration profile no longer exhibits the peak at 580 m, while following roughly the same spatial trend. However, the median is much lower than the arithmetic mean. This large difference between the mean and the median is due to the high positive skewness of the distribution. Both the trimmed mean and the COV method result in only slightly lower values than the profile based on the arithmetic mean. But the COV method misses the peak at 220 m, while with the trimmed mean this peak is still present. The application of the trimmed mean is further examined in Sections 6.3.3 and 6.3.5.

6.3.2 Spatial variation: concentration profiles

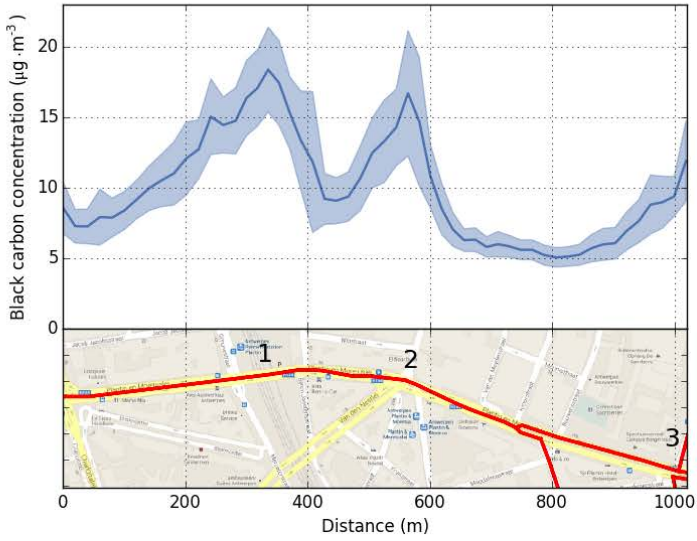
Observations on the spatial variability within the urban environment based on this dataset of mobile measurements are extensively reported in Peters et al. (2014). At the street level, significant differences in mean levels were found (up to a factor of 2) for the aggregated data of all passage means per street. A clear correspondence with traffic and street topology was observed. Lowest concentrations were found at places with low traffic intensities as Dageraadplaats and Stadspark. Highest concentrations were found in the Plantin en Moretuslei and Provinciestraat. The Plantin en Moretuslei has the highest traffic intensities, while the Provinciestraat has lower traffic intensities, but is a street canyon and has no separate biking lane.

Concentration profiles were constructed for certain sections of the routes to study the spatial variability at a high spatial resolution. The result for the 20 m resolu-

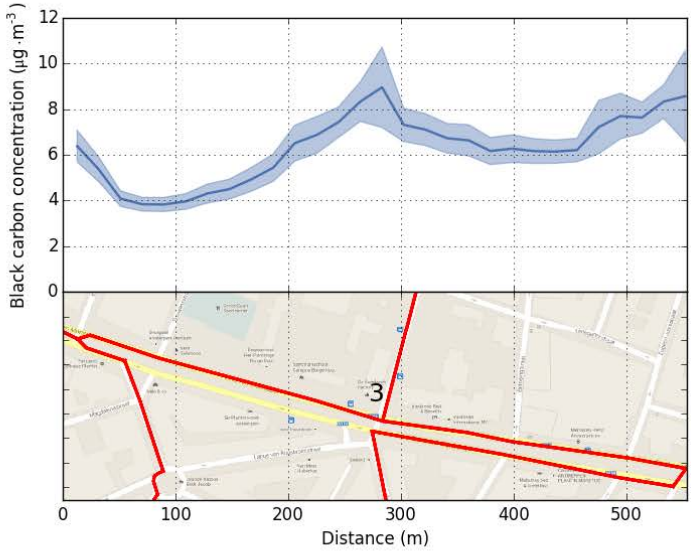
tion can be seen in Figures 6.5 and 6.6. These figures indicate that large differences in air quality exist within a single street. For example, in the Plantin en Moretuslei (Figure 6.5a) there is a very large peak at the tunnel (up to 19 $\mu g\,m^{-3}$), and a second large peak at the intersection with the Provinciestraat (up to 17 $\mu g\,m^{-3}$), where the values are two to three times higher than the BC concentrations in the rest of the street (5 to 10 $\mu g\,m^{-3}$). Also the intersection of the Plantin en Moretuslei and the Montensstraat shows an increased concentration, which is located near the VMM monitoring station (Figure 6.5b). The locations with the lowest BC concentrations are the Stadspark and Dageraadplaats (Figure 6.6, 2-3 $\mu g\,m^{-3}$). BC concentrations also show a low variability at these locations. Figure 6.6a shows a gradual increase throughout the Korte Altaarstraat to Wolfstraat from one end to the other, which is a part of route 1 from the Dageraadplaats to Plantin en Moretuslei that exhibits increasing traffic and proximity to a major traffic road. The Provinciestraat also shows a gentle gradient between two parts of the street (Figure 6.4). In contrast, a very steep gradient is observed between the Stadspark and the Quinten Matsijslei (a 2-3 times increase over a distance of 50 m, Figure 6.6b). In conclusion, large gradients over short distances between different urban micro-environments can be observed from the mobile data aggregated at resolution of 20 m . Small-scale differences up to a factor of 10 are found (in comparison to the factor 2 for street level averages). Urban greens and traffic-free squares are clearly distinguished by low BC concentrations from the surrounding streets.

6.3.3 Effect of spatial resolution on reproducibility

To investigate the reproducibility of the observed spatial variation in the concentration profiles and the influence of the spatial resolution, we look at the consistency of the spatial patterns between independent subsamples of different sizes. Two random but non-overlapping subsamples of all runs are taken. Based on each subsample a concentration profile is calculated for the whole route as the average concentration for each segment using the arithmetic or trimmed mean. The correlation between the two profiles is determined as the R^2 of a linear regression. This analysis is done for different resolutions (segment lengths between 5 and 100 m) and for different sample sizes (10, 20, 50 and 100 runs), and for each of these combinations this is repeated 1000 times to obtain an average R^2 (Figure 6.7). The average R^2 increases with increasing sample size; the concentration profiles are more similar when comparing larger samples. The R^2 also increases with increasing segment length, although this increase is large for small segments up to 20 m but almost completely levels off for segments of 40 to 50 m and larger.

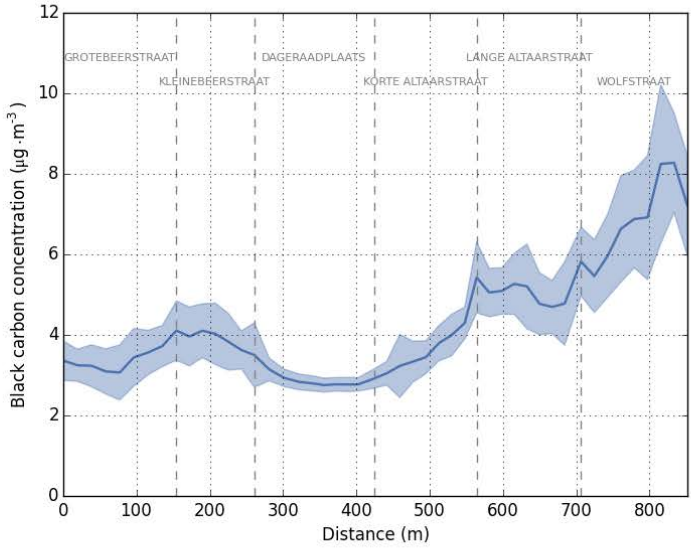


(a)

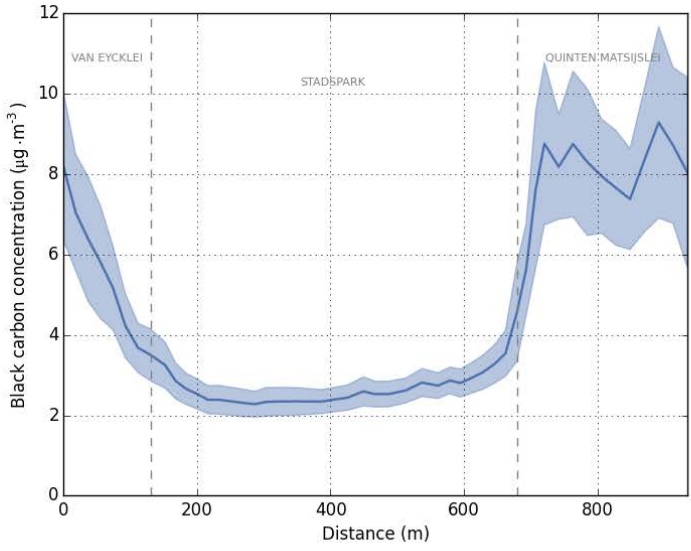


(b)

Figure 6.5: BC concentration profiles for the Plantin en Moretuslei (a) based on the measurements of route 2 and (b) of route 1 (at 20 m resolution, value is the mean of passage means). Shaded area depicts two times the standard error on the mean. The annotations indicate (1) the tunnel, (2) the intersection with the Provinciestraat and (3) the intersection with the Montensstraat and Wolfstraat.



(a)



(b)

Figure 6.6: BC concentration profiles for (a) the Dageraadplaats and (b) Stadspark (at 20 m resolution, value is the mean of passage means). Shaded area depicts two times the standard error on the mean.

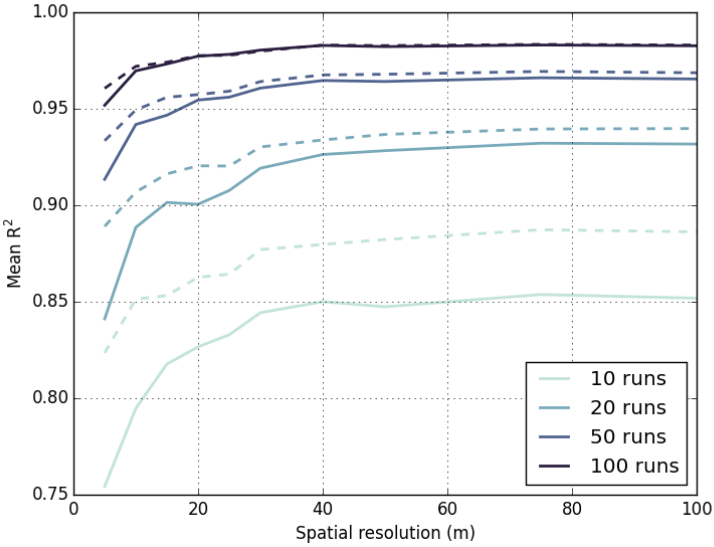


Figure 6.7: The R^2 of the comparison of two random but non-overlapping subsamples in function of the spatial resolution and for different sizes of the subsamples. The line indicates the mean R^2 of 1000 repetitions for each combination of resolution and sample size. Full lines: arithmetic mean, dashed lines: trimmed mean.

Using a trimmed mean instead of the arithmetic mean in the construction of the concentration profiles increases the R^2, and thus makes the profiles more reproducible. This increase is most clear for higher resolutions and for smaller sample sizes, while for a sample size of 100 runs it makes almost no difference.

6.3.4 Comparison of stationary and mobile measurements

Stationary and mobile measurements of BC are compared at three locations (Provinciestraat, Stadspark and Plantin en Moretuslei). For the Plantin en Moretuslei, only measurements within a range of 100 m around the stationary monitor are used because the intra-street variability is much higher than at the other two locations. Stationary measurements are available as 5 min means for each location. Mobile measurements are averaged per passage in the particular street (passage mean).

First, mobile measurements were compared directly with the stationary measurements by matching each passage mean with the nearest 5 min stationary mea-

surement. The exact time period of the measurements will not coincide using this method, and the results also indicate that there is too much variation on the individual passages to have a clear relationship between the individual mobile passages and stationary measurements ($R^2 = 0.21$, 0.40 and 0.14 for Provinciestraat, Stadspark and Plantin en Moretuslei, respectively).

When the measurements are averaged per day using the arithmetic mean, the relationship between the mobile and stationary measurements becomes stronger for all locations ($R^2 = 0.58$, 0.96 and 0.63, respectively, see Figure 6.8). For the three locations combined, this yields an R^2 of 0.69. But when we remove the one extreme value for the Provinciestraat as discussed in the second example in Section 6.3.1, the R^2 becomes 0.82. The averages of the stationary measurements are calculated based on only those measurements made during the same time frame in which mobile measurements were conducted (7 am to 1 pm). In this way, the influence of measuring during other time periods, e.g. lower night-time concentrations, which are not measured in the mobile campaign, on the mean value of the stationary measurements is cancelled out. The best correlation in both cases is found for the Stadspark, where there is less variation caused by direct sources. The combined R^2 of 0.82 indicates that the variations in time (day-to-day variations) are also captured by the mobile measurements, although the mobile measurements result in higher concentrations (positive intercept).

The same measurements are used to create a boxplot as an overall comparison for all days for the three locations (Figure 6.9). The street level averages from the mobile measurements are systematically higher than the average of stationary measurements in the same street, but this figure also indicates that the relative spatial differences between the three locations are comparable.

6.3.5 Number of repetitions based on a data experiment

A data experiment was performed as explained in the methods Section 6.2.4 in order to investigate how many mobile runs are required to obtain a representative estimate of the air quality. Firstly, the entire route is taken as the spatial level; later on the analysis is refined for smaller spatial resolutions. Figure 6.10 shows the result of one typical data experiment (based on all data of route 1 with a total of 256 runs). The evolution of the mean for all 1000 iterations is depicted on the left. For each of these iterations, the number of runs required for convergence was determined and plotted in a density plot (Figure 6.10, right), using a deviation of 25 % in this example. This number is at most 17 in 95% of the iterations. When

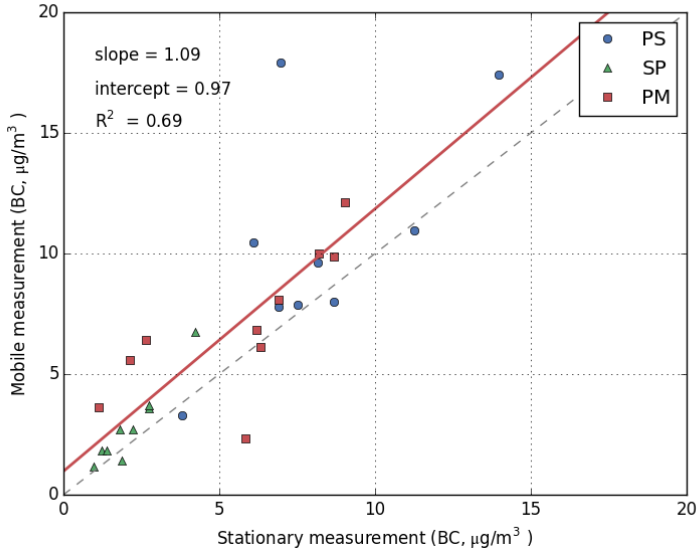


Figure 6.8: Comparison between mobile and stationary BC measurements for the three locations (for abbreviations, see Table 5.1): comparison of averages per day of passage means and stationary measurements (during the mobile time frame, i.e. 7 am to 1 pm). Grey dashed line: 1:1 line, red line: regression line.

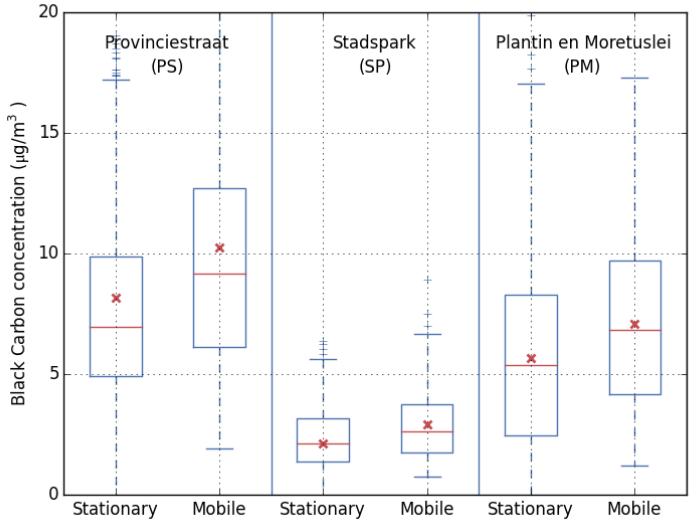


Figure 6.9: Boxplot of the stationary and mobile measurements for the three locations: comparison of the passage means (mobile data) and 5 min means (stationary data) of all days. For the stationary measurements, only the measurements during the mobile time frame (7 am to 1 pm, and also on the same days as the mobile measurements were conducted) are used.

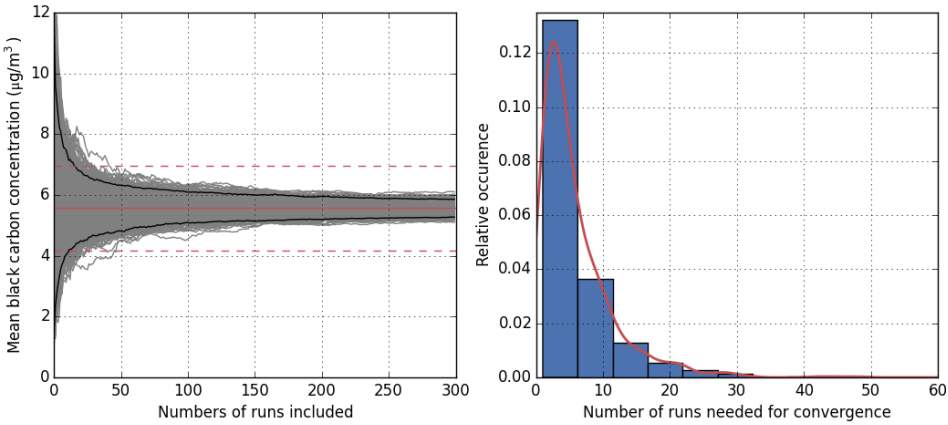


Figure 6.10: Results of a typical data experiment for all data of route 1 (total of 256 runs). Left: evolution of the mean value in function of the included number of runs for 1000 iterations. The overall mean and the deviation of 25 % around this mean are indicated by the horizontal red lines. The black lines indicate the 2.5 % and 97.5 % percentiles of the mean value. Right: density plot of the required number of runs to obtain convergence for these iterations.

narrowing the deviation to 20 %, this number increases to 29 runs, and it drops to 12 runs when allowing a deviation of 30 %.

This data experiment is repeated for each street of route 1 separately. Due to different pollution dynamics, different results are expected in the different streets. The results are summarised in Table 6.1 for some streets of route 1 for a deviation of 25 and 50 %. In general, the required number of repetitions for the separate streets is higher than for the overall mean. This is caused by the increasing variation between the runs when going to a higher spatial resolution because the sample becomes smaller per spatial entity. The mean values at Dageraadplaats and Plantin en Moretuslei converge more quickly. At the traffic-free square of Dageraadplaats, there is less variation between different runs, which explains the lower number of repetitions required to obtain convergence. The Plantin en Moretuslei is a busy street and a large variability is expected, but it is also the longest street, levelling off this variability by a larger temporal smoothing. On the other hand, the mean value at the Wolfstraat and Kleinebeerstraat converges slowly. These streets are both shorter and the Wolfstraat ends at the crossing with a busy street. This illustrates a shortcoming of using street sections with varying length as the spatial aggregation level. Therefore, the same analysis is performed with

Table 6.1: Results of the data experiments: required number of repetitions using a deviation of 25 and 50 % at different levels of spatial aggregation: the entire route (route 1), different streets (for abbreviations see Table 5.1) and 20 and 50 m segments (10th, 50th and 90th percentiles are given), and using different processing methods: arithmetic mean ('standard'), trimmed mean ('trimmed'), with a background normalisation ('background') and the combination of both ('both'). The minimum for each spatial level and for 25 and 50 % is shown in italics.

		25 %				50 %			
		standard	trimmed	background	both	standard	trimmed	background	both
Route		17	18	*11*	12	4	5	*3*	4
Street level									
	PM	18	20	*13*	14	5	6	*3*	4
	DP	18	19	*14*	*14*	5	5	*4*	*4*
	WO	63	61	60	*57*	16	11	15	*9*
	KB	82	68	75	*61*	21	10	20	*9*
	GB	66	50	62	*42*	18	8	18	*7*
	KA	44	34	40	*25*	12	7	12	*5*
20 m segments									
	10 %	31	30	26	*24*	8	6	7	*5*
	50 %	59	54	52	*48*	15	10	13	*8*
	90 %	164	112	156	*108*	44	13	43	*12*
50m segments									
	10 %	33	31	29	*24*	8	7	6	*5*
	50 %	57	52	50	*42*	14	10	13	*8*
	90 %	141	102	141	*94*	39	13	40	*11*

segments at 20 m and 50 m resolution. There is a large variation in the number of repetitions required for the different segments. In Table 6.1, this variation is expressed as the 10th, 50th and 90th percentile of the required number of repetitions for the different segments. In 10 % of the segments convergence occurs after less than 33 repetitions, while in 10 % of the segments this is only after more than 141 runs at a spatial resolution of 50 m. In general the required number of repetitions is only slightly lower at 50 m resolution compared with the higher resolution of 20 m.

Next, a similar data experiment is carried out with a trimmed mean instead of the arithmetic mean (Table 6.1). On the whole dataset, this does not have a positive effect, but when looking at specific streets, this can have an impact. For example, for the Grotebeerstraat (street with an outlier), the required number of repetitions decreases from 66 to 50 and for the Kleinebeerstraat from 82 to 68 for a deviation of 25 %. Using a deviation of 50 % the relative decreases in the required number of repetitions are larger. In some cases, however, there is no or even a negative effect. For the Plantin en Moretuslei and Dageraadplaats, the required number of runs increases.

A third experiment is carried out to investigate the use of background normalisation. Background normalisation decreases the variability between runs by accounting for temporal variations in the background concentration, and consequently also the required number of repetitions. The background normalisation clearly decreases the required number of repetitions for all streets compared with the basic experiment, up to 30 % for Dageraadplaats and Plantin en Moretuslei. The background normalisation was performed with both types of background concentrations (hourly median and moving mean), but no significant difference was noticed in the resulting required number of repetitions. Therefore, only the results obtained with the hourly median background concentration are shown in Table 6.1.

The lowest values for the required number of repetitions are found when combining the use of the trimmed mean and performing the background normalisation. Using this combination reduces the required number of repetitions with 25-33 % (35 - 70 % for 50 % deviation) compared with the initial data experiment. This leads to a number of 24 to 94 repeated measurements runs (10th and 90th percentiles, median of 41) for the different 50 m segments with a deviation of 25 %. When relaxing the deviation to 50 %, these numbers reduce to 5 to 11 (median of 8) runs.

6.4 Discussion

6.4.1 Mobile monitoring to assess spatial variation

The BC concentrations in the urban environment show a high variability due to differences in, amongst other, traffic intensity, street topology and proximity of other sources. The results of this case study confirm this variability, showing large variations even over short distances within a single street (see concentration profiles in Figures 6.5-6.6). The measured average concentrations at 20 m resolution range from 2-3 $\mu g\,m^{-3}$ in the Stadspark to values of 19 $\mu g\,m^{-3}$ at the tunnel in Plantin en Moretuslei, while an average concentration of around 6 $\mu g\,m^{-3}$ is measured at the monitoring station.

It is clear that stationary monitors are not very well suited to study this level of variability, and that the result can vary greatly depending on where in the street the monitor is placed. This strongly confirms that stationary monitoring stations are not always representative for the population exposure (Kaur et al., 2007). Mobile monitoring, on the other hand, is a suitable monitoring approach for mapping the air quality at a high spatial resolution, as shown with the results above. It can provide insight into the spatial variability that would not be possible with stationary monitors. Previous case studies already showed the potential of repeated mobile measurements for the assessment of spatial variability of pollutants at different micro-environments in a city at street level (Van Poppel et al., 2013; Peters et al., 2013b) and within street-level (Zwack et al., 2011b, Chapter 5).

6.4.2 Influence of events on spatial concentration patterns

The concentration profiles discussed in the previous section show a large variability of the BC concentrations in space, but from the time series on Figure 6.2a it is clear that there is an even larger variation in time when looking at the one second resolution measurements. With mobile monitoring, measurements at different locations at different times are obtained. As a consequence of this moving monitor, the short-term variability in time can cause an apparent variation in space. However, this variation should not necessarily be attributed to that specific location. The magnitude of this impact and possible solutions to deal with it is investigated.

The results in Section 6.3.1 clearly show a potentially large influence of extreme events on the average measured BC concentration both at the street level as for segments, and thus on the resulting spatial pattern. This impact is relatively small

for all 256 or 96 runs (for route 1 and 2, respectively), but for samples of more realistic size it is much more prominent. For the examples illustrated in Section 6.3.1 and with a sample size of 20 runs, the influence of one run with an extreme event can go up to 50 to 250 % for a street and 20 m segment, respectively.

To minimize the influence of the short-term temporal variation, data are smoothed, i.e. several one second measurements are averaged over time. In this study, a spatial smoothing is applied by using passage means at different spatial levels. But even at the street level, events can still have a large impact. The initial spatial smoothing is thus not sufficient to deal with all of the short-term temporal variation at this spatial resolution. Therefore, in a second step, filtering methods or alternative statistics for the average are applied.

Before looking into the processing and possible removal of peaks, it is important to ask the question whether they are relevant to the aim of the measurements. Depending on the application, it can be sensible or not to remove some of the extreme events. For example, in the study of Hagler et al. (2012), an event detection algorithm is used to remove these extreme values from the dataset. In their study, the purpose was to measure the concentration profile in a near-road environment to investigate the influence of a highway. Therefore, a passing car could distort the result heavily as they are less interested in the local emissions of that car. Hudda et al. (2014) and Choi et al. (2012) used a rolling percentile to obtain a baseline concentration to discard the local peaks, as they wanted to map the larger scale variations due to the impact of an airport and a highway, respectively. However, in an urban environment with busy traffic, peak concentrations are an integral part of the air quality that people are typically exposed to. If the goal of the measurement campaign is to map the local concentrations to which cyclists are exposed, it is important to include the peak concentrations. On the other hand, as we have seen, the rare occurrence of large peaks can also distort the concentration profile to a large degree.

To deal with these events and to reduce their impact, we looked at filtering methods and alternative statistics for the arithmetic mean. The median can be used as a more representative central tendency measure as opposed to the mean since air quality measurements are not normally distributed and skewed, and so the median tends to be more robust to bias caused by emission events (Peters et al., 2013b; Brantley et al., 2014b). However, the results show that the median also leads to much lower BC concentrations than the average concentration. Being interested in the average concentration, the median will not be used further in this study. A trimmed mean, with trimming performed for each spatial entity separately, was used instead of the median. The trimmed mean appears to be less sensitive to

peaks than the arithmetic mean, but without the large underestimation of the average concentration compared to the median. The COV method as used in Hagler et al. (2012) and Brantley et al. (2014b) to detect events also shows good results. However, this method has the risk that it can mask hotspots where peaks occur systematically. This is also reflected in the results, where the peak at 225 m is masked when using the COV method. Recurring peaks at the same place should not be removed but accounted for because they are relevant. Therefore the COV method is not appropriate for the goal in this study and is not used further, although in cases such as Hagler et al. (2012) where the goal is not to measure the sources in direct vicinity, this can certainly be a valuable method. The trimmed mean does not have this issue: when a peak systematically occurs at a certain location, the trimmed mean will only remove part of it and the peak will still be present in the concentration profile. When having smaller sample sizes, the risk of removing an extreme event that was not random but systematic becomes larger.

The results discussed above show that the arithmetic mean can potentially exhibit a large bias due to the occurrence of extreme events. The trimmed mean can be used to counter this, but it is delicate to draw a line between the detrimental impact of an event and typical peak exposure at a location. The trimmed mean seems to hold a good trade-off between these two: it still measures systematically occurring peaks leading to a typical concentration profile, but reduces the detrimental impact of random events on the result. But most important is to be aware of this issue when performing and analysing this kind of measurement campaigns, and taking appropriate actions such as looking at different statistics and investigating locations were large differences between those statistics appear, in more detail. The application of the trimmed mean will further be validated in the following sections on spatial resolution and number of repetitions.

6.4.3 Spatial resolution

This study deals with mapping the air quality at different levels of spatial data aggregation (route, streets, segments). But which resolution is relevant and achievable? The entire route can represent the mean exposure of a cyclist on that trajectory, or an estimation of the average concentration in a neighbourhood. The street level gives more detail. This level can be a somewhat artificial delineation, as there are streets with different length and large variations within a single street exist (e.g. the Plantin en Moretuslei, Figure 6.5), but can be easy to communicate. The segments of certain length provide a more systematic and higher spatial resolution. From Figures 6.5 and 6.6 it is clear that the urban air quality (BC levels)

can exhibit large variations over short distances. Differences up to an order of 3 over distances of around 50 m are seen. To be able to map the urban air quality and identify hotspots, higher resolutions up to 20 - 50 m are useful. Dependent on the application, such a high resolution can be relevant.

Apart from the relevance to measure air quality at such a high resolution, it is also important to see whether it is achievable with this type of monitoring. The spatial accuracy of the GPS is estimated at 10 m, so a resolution below 10 m is not feasible. The large influence of extreme events on the concentration profile has been discussed before. The question can be raised if it is not the case that a large part of the spatial variation in the profile is caused by such random events. The analysis in Section 6.3.3 (Figure 6.7), however, indicates that mapping the air quality at a resolution of 40-50 m leads to consistent profiles and is reproducible. Using higher spatial resolutions up to 20 m is also possible with a small increase in the uncertainty. Of course, this depends on the sample size as the concentration profile will be more accurate or a higher spatial resolution can be used with more repeated runs. Enlarging the grid size from 50 to 100 m does not improve the reproducibility of the profile significantly. When using larger segments, data are smoothed over a longer time period and a reduced variability from events is expected. But at the same time, this will also potentially combine street sections that can be very different, and as such increase the heterogeneity within a single segment. For smaller sample sizes, it is more important to deal with events (in this case using the trimmed mean).

6.4.4 Temporal representativeness

For a particular location, mobile monitoring gives only a very sparse coverage in time compared with a continuously measuring stationary monitor. This raises the question whether a limited mobile dataset, sparse in time, can be used to estimate the average concentration for the full time period during which the mobile measurements are carried out (in this case 7 am - 1 pm during 4 weeks). To validate the mobile measurements, they are compared with stationary measurements. Because the measurements are not performed at exactly the same place and with the same time resolution, and only three locations are available, the strength of the comparison is limited. But when it is assumed that the stationary monitor is representative for the neighbouring, relatively homogeneous street section, comparing the mobile and stationary measurements can give a valuable indication of the representativeness in time of the mobile measurements.

The results of this comparison indicate that with a limited set of mobile measurements it is possible to find indicative results similar to stationary measurements. The comparison of the average values per day for the same time period (Figure 6.8) shows that the variations in time (day-to-day variations) are also captured by the mobile measurements. From Figure 6.9 it can be concluded that the mobile measurements are able to map similar spatial differences as a stationary monitor at each location, although the mobile measurements are consistently slightly higher than the stationary measurements. This can be explained by the shorter distance to the traffic compared with the stationary monitoring locations.

Despite the limited temporal resolution at one point in space, it can be concluded that mobile measurements are able to assess the air quality at a fixed location as could be performed with a stationary monitor. The comparison is performed for three locations, but it is assumed this conclusion can be extrapolated to other locations where mobile measurements were done. Mobile measurements give a representative image of the local concentrations (with certain spatial aggregation level, see previous section), and in this way mobile monitoring can provide for a spatial coverage that is not possible with stationary monitors.

The analysis here is limited to the time period (hours and days) of the mobile measurement campaign, but the question on representativeness could be broadened to longer periods. This could possibly be done with a background normalisation to adjust the measurements for temporal variation, as is also done for short-time stationary measurement campaigns (e.g. Hoek et al., 2002; Eeftens et al., 2012b). However, this falls outside the scope of the present study.

6.4.5 Number of repetitions

The case study presented in this study, consisting of 256 runs of a specific route, is a very extensive monitoring campaign. In the previous section, the analysis was based on all 256 or 96 runs (for route 1 and 2, respectively). Here, we will discuss whether it is possible to obtain this representative image of the local concentrations with a smaller number of runs. The overall mean of all runs is used as a good estimate of the 'true' average concentration at that place.

The results of the data experiment summarised in Table 6.1 indicate clearly that it is indeed possible to achieve similar results with less repetitions. The required number of repetitions depends on the desired confidence of the result. When decreasing the deviation from the overall mean from 50 % to 25 %, the required number of repetitions increases with a factor of 3 to 9 depending on the location.

The required number of repetitions does also strongly depend on the location, as the variation of the measurements at each location is different.

Using the trimmed mean decreases the number of runs up to 33 % for certain locations (for a deviation of 25 %). So by excluding part of the extreme measurements, the trimmed mean converges more rapidly to the overall mean based on all measurements (computed without trimming) compared with the arithmetic mean. As such, we get in general a better estimation of the average BC concentration with the trimmed mean. However, for other streets, it leads to a slight increase in the required number of repetitions. This is due to the fact that using the trimmed mean introduces a systematic bias from the 'real' mean, and for locations with a limited number of extreme values this will lead to an increase of the required number of repetitions. But in general, it has a positive effect and the combination of the trimmed mean and a background normalisation does lead to a decrease in the required number of repetitions for all locations. So the trimmed mean and background normalisation are valuable approaches to reduce the required number of repetitions or improve the reliability for a given number of repetitions. The results indicate that it does not matter which of both types of background concentration values is used. Including temporal variation in the background concentration at below-hourly scale by using a smoothed moving mean did not further reduce the required number of repetitions compared to using hourly medians in the background normalisation.

The numbers of 25 and 50 % used for the deviation give an indication of the uncertainty of the resulting average concentration. They denote that, if a mobile campaign is conducted, the resulting concentration will be within 25 or 50 % from the 'true' average. The European directive on ambient air quality (2008/50/EC) uses these numbers of 25 and 50 % as the data quality objectives for particulate matter for fixed and indicative measurements, respectively. However, note that this deviation only includes the uncertainty coming from the variability in the measurand noticed by repeated measurements, and does not include other sources that contribute to the data quality (e.g. the uncertainty on the instrument).

A comparable analysis with this data experiment was performed in Peters et al. (2013b) and Van Poppel et al. (2013). They concluded that a limited set of mobile measurements (20 to 24 runs) makes it possible to map the air quality, and for some streets this number could even be reduced to less than 10 (Peters et al., 2013b). They used the median as the central tendency measure and a deviation of 15 %. Van Poppel et al. (2013) showed that the use of background normalisation leads to a faster convergence towards a representative concentration and reduces the number of runs required to obtain representative results significantly,

which is in line with the results in the present study. In those papers, a more limited number of repetitions to perform the data experiment was used. While their conclusions on the variation between streets and background normalisation are still valid and are confirmed here, the results in the present study contradict their conclusion on absolute numbers of required repetitions. This is likely due to the limited number of runs in their initial dataset and the different sampling method in the data experiment (sampling without replacement instead of with replacement). Xu et al. (2007) also use a similar approach but to develop another sampling strategy, namely repeated short-term stationary measurements with a mobile laboratory.

The results of the data experiments are limited to the constraints of the available dataset. Firstly, the analysis is constrained by the space and time of the data collection, which was biased in this study towards daytime hours during one month of the year in an urban environment, and hence is not necessarily representative for other periods or locations. The data collection mainly focused on the time of the day when the majority of the people are exposed to outdoor pollution, and did not include night-time. The measurements were equally spread over 11 days including a wide range of background concentrations (meteorological conditions). Secondly, one of the disadvantages of the approach of the data experiment is the dependence on the size of the dataset. The dataset has to be large enough to be able to draw justified conclusions. But given that the results for the required number of runs shown in Table 6.1 are well below the total number of 256 runs, we conclude that the number in this case study is enough. Lastly, in absence of continuous measurements, the overall mean of the dataset is used as an approximation of the pollutant concentration to be measured to determine convergence, as the 'real' concentration we want to measure is not known. The large number of measurements and the preceding analyses in Section 6.4.4 on the comparison of stationary and mobile measurements indicate that this approach is justified and solid conclusions can be drawn.

6.5 Conclusion

Mobile monitoring is a suitable monitoring approach for mapping the urban air quality at a high spatial resolution. Mapping at a spatial resolution up to 20 m is relevant as large gradients over short distances are observed and differences up to a factor of 10 in BC concentrations are measured at this resolution. As we have shown in this chapter, mapping at such a high resolution is also possible, but a lot of repeated measurements are required. This can, however, be reduced by dealing

with extreme events using the trimmed mean and by performing a background normalisation. Based on a targeted measurement campaign with 256 runs and after computing a trimmed mean and applying background normalisation, we found that in this case study 24 to 94 repeated measurement runs (median of 41) for the different locations are required to map the BC concentrations at 50 m resolution with an uncertainty of 25 %. When relaxing the uncertainty to 50 %, these numbers reduce to 5 to 11 (median of 8) runs. Given the increasing use of mobile air quality monitoring and its potential for participatory sensing and crowd-sourcing projects, the representativeness of the data that are collected in this way is a key issue. A careful set-up is needed with a sufficient number of repetitions in relation to the desired reliability and spatial resolution, dependent on the aim of the campaign. Specific data processing methods such as background normalisation and event detection are applied. Given this careful set-up and processing, mobile monitoring can provide insight into the spatial variability that would not be possible with stationary monitors.

CHAPTER 7

Opportunistic mobile air pollution monitoring: a case study with city wardens in Antwerp

7.1 Introduction

As shown in the previous chapter, to map the urban air quality in a reproducible way and at a high spatial resolution, a sufficient number of repeated measurements is required. However, this is generally much more labour intensive compared to stationary measurements. One way to collect the large amounts of data that are needed, is to take advantage of existing mobile infrastructure or people's common daily routines to move measurement devices around through the city, without specifically designing the travelled route of the mobile carrier for the measurement campaign. We call this *opportunistic mobile monitoring*, as explained in Chapter 4.

In this chapter, we describe a case study using such an opportunistic data collection scheme, consisting of mobile measurements of black carbon (BC) and carried out with the collaboration of city wardens of the city of Antwerp. As a consequence of their working routines, there is no control over the followed route and the campaign has an unstructured set-up. The data collection is unevenly spread in space and time, leading to sampling bias as different locations will possibly be measured at

This chapter is based on the following publication: Van den Bossche, J., Theunis, J., Elen, B., Peters, J., Botteldooren, D., and De Baets, B. (2016). Opportunistic mobile air pollution monitoring: a case study with city wardens in Antwerp. *Atmospheric Environment*, 141 408-421.

different days and/or hours of the day and complicating the data interpretation. This is a major problem which has to be addressed to be able to compare the results between these locations and to use the results for air quality mapping. The issues discussed in the previous chapter concerning the representativeness of mobile measurements, the sensitivity to peaks and varying background concentrations, will now be even more important, and will have to be re-evaluated.

The goal of this chapter is to explore the potential of an opportunistic mobile monitoring approach with an unstructured set-up, as could for instance be the case in an unstructured participatory monitoring campaign, to obtain a reliable, high-resolution map of the urban air quality (not a city-wide map but restricted to the measured locations). It will be evaluated whether or not the conclusions for mapping the urban air quality with a structured mobile monitoring campaign from the previous chapter can be applied here. The focus of this chapter is on the characterization of the sampling bias, its impact on the result and ways to counteract the sampling bias through the use of background normalisation. The resulting air pollution concentrations will be compared to concentrations from a targeted campaign.

7.2 Materials and methods

7.2.1 Opportunistic measurement campaign with city wardens

A case study was set up to map the average exposure to black carbon (BC) in streets and public spaces in Antwerp (51°12' N, 4°26' E, medium-sized city of 480,000 inhabitants, 985 inhabitants km^{-2}), Belgium, during a one-year period (from July 2012 until June 2013). The case study was carried out with the collaboration of city wardens and the Environmental Services of the city of Antwerp. The Antwerp city wardens are city employees with a surveillance task. They are outdoors for a large part of the day carrying out surveillance tours by bicycle or on foot. These surveillance tours do not follow fixed routes or times. Three teams of 2 city wardens each were equipped with a measurement unit, using the VITO airQmap platform[1] (see Figure 7.1a). The measurement unit consists of a micro-aethalometer (MicroAeth Model AE51, AethLabs), a lightweight sensor that allows to measure BC at a high (1 s) frequency, and a GPS (Locosys Genie GT-31 GPS). The airQmap platform further consists of a home station, a netbook

[1] http://www.airqmap.com

(a) (b)

Figure 7.1: (a) Micro-aethalometer and a GPS together in a belt bag as used in airQmap and (b) members of the 3 teams of city wardens measuring air quality in Antwerp during 1 year (© De Nieuwe Antwerpenaar).

with custom-made, easy to use software to read out the measurement devices, keep their clocks synchronized and to transmit the measurements over the internet. The city wardens were asked to read out and charge the measurement devices on a daily basis, and to replace the filters of the micro-aethalometers every second day. There was limited follow-up apart from technical support to the instrumentation. Because of the risk of damage to the micro-aethalometer, the city wardens were asked not to take measurements when it was raining. Each measurement is automatically linked to its geographic location by time-synchronizing the BC and GPS data. Further processing steps including the filtering for unreliable GPS data, BC noise reduction and calibration and spatial aggregation are described in more detail below. The study area (i.e. the surveillance area of the three teams together) covers approximately 3.6 km^2 (a quarter of the inner city), and is pictured in Figure 7.2.

During the measurement campaign, 393 hours of raw 1 second measurements were recorded for the three teams combined (459 hours of measurements before filtering for GPS quality), spread over 110 days. Most of the measurements were done between 10 am and 16 pm during working days.

Additional stationary measurements from the official monitoring stations of VMM (Flemish Environmental Agency) were available. At these stations, BC is measured with Multi-Angle Absorption Photometry instruments (MAAP, Petzold and Schonlinner, 2004; Petzold et al., 2005. One location with two monitoring stations,

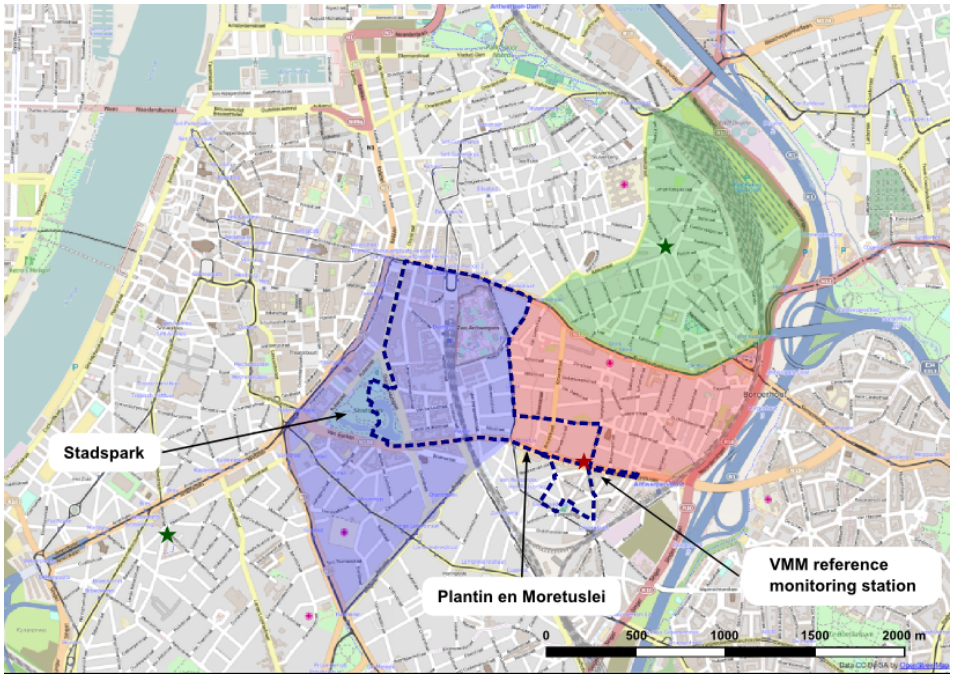


Figure 7.2: Overview of the study area. The three coloured areas indicate the zones of the three teams of city wardens and the two green stars indicate their offices. The blue dashed line indicates the route taken in the targeted campaign described in Chapters 5 and 6. The red star indicates the location of the official monitoring station.

BETR802 and BETR801 (at the street side and 30 m away from traffic, respectively) is situated in the study area (red star in Figure 7.2).

7.2.2 Quality control and processing of the BC data

Measurements with the micro-aethalometer were made at a temporal resolution of one second, and the inlet flow rate was set at 100 mL min^{-1}. To reduce the signal noize, the ONA algorithm was used with an attenuation threshold of 0.05 (Hagler et al., 2011). To limit the filter loading effect, the filter ticket was changed every 2 measurement days, ensuring that 95 % of all measurements have an ATN value below 100. To account for the remaining filter loading effects, the correction algorithm according to Virkkula et al. (2007) was applied. The parameter k of the correction algorithm is estimated based on the comparison to the MAAP instrument from the monitoring station. More details on the processing of the micro-aethalometer data can be found in Chapter 3.

Different micro-aethalometers were used throughout the campaign (different instruments for each team, and different instruments were used in the first half and second half of the campaign). For data quality control, the individual micro-aethalometers were all deployed at the VMM monitoring station at the Plantin en Moretuslei (station BETR802) to identify possible bias between instruments. After the filter loading correction and comparing to the mean of all instruments, the micro-aethalometers showed on average a mean absolute error of 0.4 $\mu\mathrm{g\,m}^{-3}$ at 1 minute resolution and 0.1 $\mu\mathrm{g\,m}^{-3}$ at 30 minute resolution (15 % and 5 % error, respectively). When comparing the micro-aethalometers to the MAAP of the reference station, the overall bias ranged between 0.04 and 0.27 $\mu\mathrm{g\,m}^{-3}$ (bias of 1 to 4 %), and a mean absolute error of 0.4 $\mu\mathrm{g\,m}^{-3}$ was found at a 1 hour resolution (15 % error).

7.2.3 Quality control and processing of the GPS data

The raw measurements of the GPS device are filtered and processed to ensure data quality. The different processing steps consist of filtering for incorrect or unreliable GPS locations, map matching and spatial aggregation. Firstly, the GPS data are filtered by quality. Of all available GPS data, 71 % had a fix (i.e. a minimum of three satellite signals to calculate the position); the remaining data probably consist of indoor periods during the day or outdoor moments with a very bad reception. Starting from these data, some filtering criteria were adopted to remove incorrect or unreliable GPS locations: 1) assuming all data were gathered

by bike or on foot, the speed may not be above 25 $km\,h^{-1}$ and 2) to ensure the GPS quality, the 30 s centred moving mean of the number of satellites should be above 4. After these filtering steps, 86 % of the data with a fix remained, corresponding to 393 h of measurements.

Secondly, map matching was applied. In urban environments, the location determined by the GPS is often slightly off track. A commonly used approach to enhance the obtained GPS tracks is map matching, assuming the measurements are always performed on the streets. For each measurement day, those street sections are selected from the database at which measurements have been conducted over the entire section. In a next step, a simple map matching based on shortest distance (max 30 m) is applied for each measurement day separately with the selected street sections. This selection of streets per measurement day is done to exclude map matching to (side) streets that were not actually visited. During map matching, an additional 4 % of the filtered GPS data is removed as they are not within 30 m of a street section. In the supplementary data (Section 7.A.1), some GPS tracks are included as an illustration of the raw and map matched tracks.

7.2.4 Data analysis: spatial aggregation, trimming and background normalization

For further analysis of the mobile measurements, the data are spatially aggregated. For this purpose, the street sections are divided into segments of approximately 50 m. All measurements are allocated to such a segment. For each passage through a segment (*passage* = continuous period of time measurements are performed in one segment), an average BC value is calculated for the segment (= *passage mean*). To assure data quality, we require the calculated average to be based on at least 8 one second measurements (which corresponds to covering 50 m at an assumed maximum speed of 25 $km\,h^{-1}$). This removes an additional 3.5 % of the data and leads to 40,284 recorded passages. The passage means are further averaged to obtain the mean concentration for each segment (= *segment mean*).

To reduce the impact of extreme values on the average concentration of a segment, a trimmed mean is applied as described in Chapter 6. The mobile measurement data are normalized with a background normalization method to minimize the influence of meteorological day-to-day variations in the urban air quality (see Eq. (6.1)). Background concentrations are obtained from the stationary measurements at VMM stations in the Antwerp region (mean of the stations BELAL01, BETM802, BETN016, BETR801 and BETR802).

The statistical comparison of the pollutant concentrations for cycling and walking data is based on the Kruskal-Wallis test at a critical value of 0.05 (Kruskal and Wallis, 1952). The Kruskal-Wallis test is a non-parametric alternative to a one-way ANOVA and investigates whether the measurements come from the same population.

Analyses were performed using Python pandas (PyData Development Team, 2014) and QGIS (QGIS Development Team, 2014).

7.3 Results

7.3.1 Characterization of sampling bias

Spatial coverage

Due to the nature of their working routines, the city wardens pass by at varying locations in the city. After one year of measurements, this resulted in a high spatial coverage. Figure 7.3 shows a map with the number of times they passed by at a certain segment. Almost all street segments were covered at least once within the study area (see Figure 7.2), and 74 % of the segments were measured at least 5 times. Only those segments are shown in Figure 7.3. The number of measurements varies greatly in space. Some segments were measured more than 400 times, typically those locations close to the wardens' office (start and end point of the measurements). Most segments, however, were measured 9 to 27 times (interquartile range).

Temporal coverage

Apart from the varying spatial coverage among street segments, also their temporal coverage varies. When looking at all data together, the temporal coverage (distribution of the number of passages in time over the day and over the year) is not uniform (Figure 7.4). More measurements were performed at the beginning and the end of the measurement campaign, resulting in more measurements during the summer compared to the winter months. Further, during the day, most measurements were performed between 9 and 16 h. Combined with varying typical concentrations in these periods of the year and the day (illustrated with the concentrations of the official monitoring station that lies within the study area, Figure 7.6), this will introduce a bias in the results. The winter months typically have higher concentrations, but fewer measurements are conducted during this

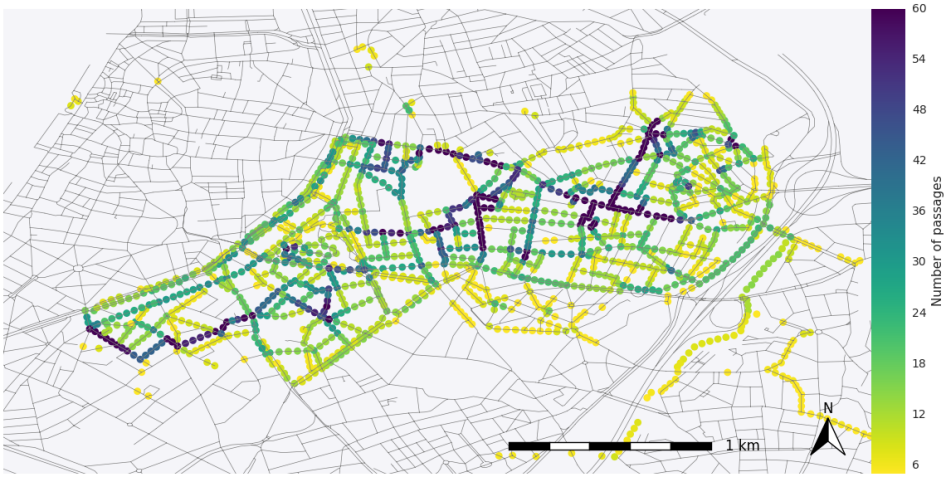


Figure 7.3: Number of passages at each segment. Segments with less than 5 passages are excluded, and the colour scale is truncated at 60 passages (some segments near the office of the city wardens have more than 400 passages).

period. The same applies to the morning and evening rush hours. The morning rush is not covered by the measurements, resulting in lower concentrations. Furthermore, this temporal coverage also varies between locations. Examples are given for the Stadspark and Provinciestraat (Figure 7.5), where in the Provinciestraat relatively much more measurements were collected in the morning hours. This varying temporal coverage between locations will introduce a different bias between segments.

Mode of transport: walking and cycling

Another possible source of bias in the data set is the mode of transport. The data are gathered both by bike and on foot. It is possible that the cyclists are exposed to higher concentration levels compared to the pedestrians (as being in general closer to the traffic). Next to the difference in temporal and/or spatial coverage, this can be an additional source of variation in the measured concentration levels when data of the two modes of transport are mixed.

The data set is divided by mode of transport (on foot or by bike). Based on the speed measurements of the GPS, a simple division is made based on a cut-off of 5 km/h on the 1 min rolling average speed. The data with low speed are classified as walking data, and those with a speed above 5 km/h as cycling data. Short stops

Figure 7.4: Overview of the temporal distribution of the number of passages in time. The horizontal histogram indicates the distribution over the year (per month) and the vertical histogram the distribution over the day (per hour). The internal plot is a combination of both and indicates the number of passages for each hour and each week (the legend for this figure is the colour scale on the right).

during cycling are dealt with by using a rolling average of the speed. Of all second measurements, 25 % are classified as cycling data. Both subsets are processed in the same way as described in Section 7.2.4 (by taking passage means).

The measured BC concentrations while cycling have a clearly higher overall average: 4.1 vs 2.6 $\mu g\,m^{-3}$ for cycling and walking data, respectively. However, the measurements are not necessarily made at the same locations. When using only those segments that have data for both modes, the difference decreases (averages of 3.6 vs 3.0 $\mu g\,m^{-3}$, respectively). Figure 7.7 also illustrates this difference in a histogram of both data sets. The cycling data are slightly shifted towards higher values. Based on the non-parametric Kruskal-Wallis test, this difference is significant ($P < 0.05$). Given the distribution of the passage means (Figure 7.7), the difference of 0.6 $\mu g\,m^{-3}$ is, however, not large; it is smaller than the standard deviation in passage means (1.8 and 1.5 $\mu g\,m^{-3}$ for cycling and walking data, respectively).

The results above indicate that there is an observable difference in the concentration levels of cycling and walking data. This can be a systematic difference, due to a higher exposure of cyclists. But first some other factors have to be eliminated. Another possible explanation of the difference is the time period of the measurements. The distribution over the year is not notably different between

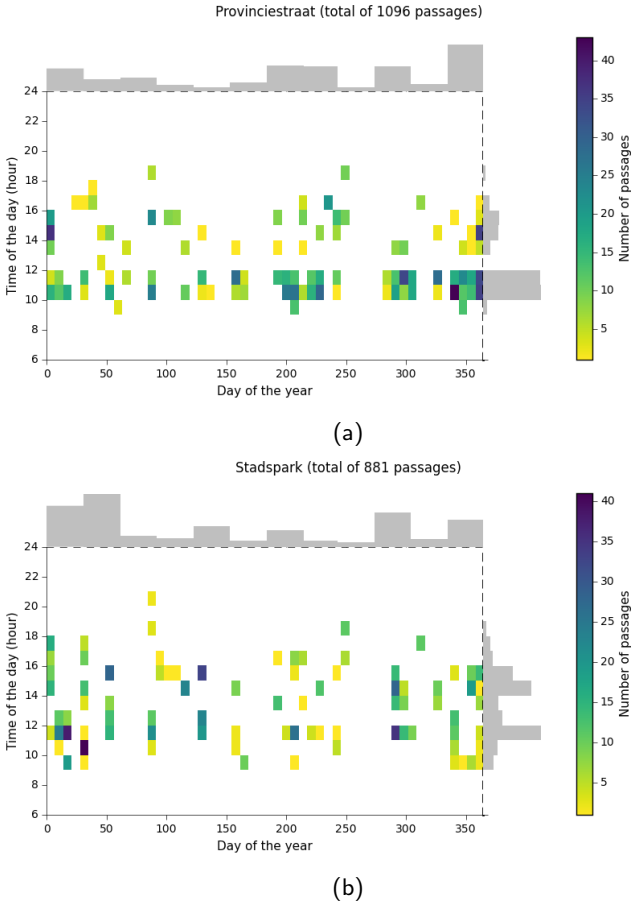
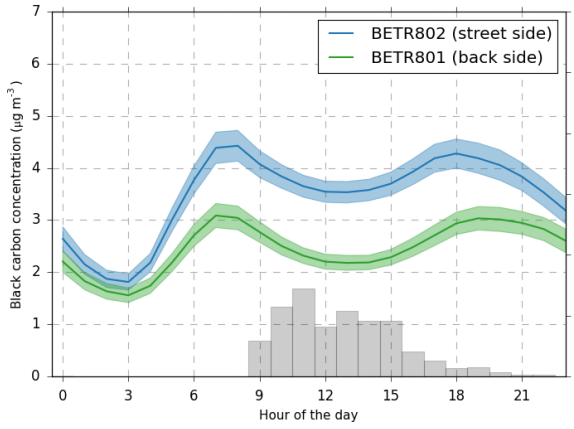
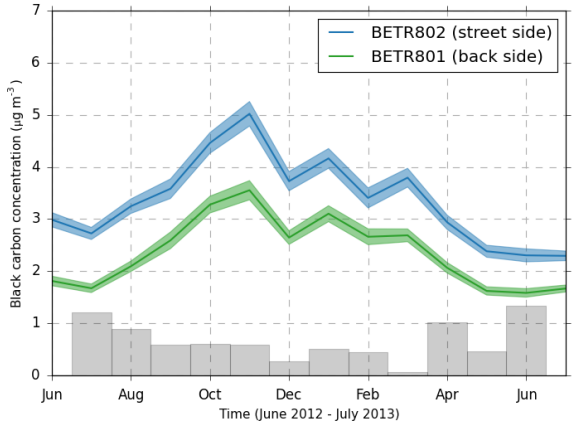


Figure 7.5: Overview of the temporal distribution for two different locations: (a) Provinciestraat and (b) Stadspark. The horizontal histogram indicates the distribution over the year (per month) and the vertical histogram the distribution over the day (per hour). The internal plot is a combination of both and indicates the number of passages for each hour and each week (the legend for this figure is the colour scale on the right).



(a)



(b)

Figure 7.6: BC concentration at the reference monitoring station near the study area (for two nearby stations, BETR801 and BETR802). Typical daily profile of the concentrations (hourly values averaged over all days during time frame of the campaign) is given in (a) and the monthly averages during the time frame of the campaign in (b). The histogram at the bottom indicates the relative distribution of the mobile measurements.

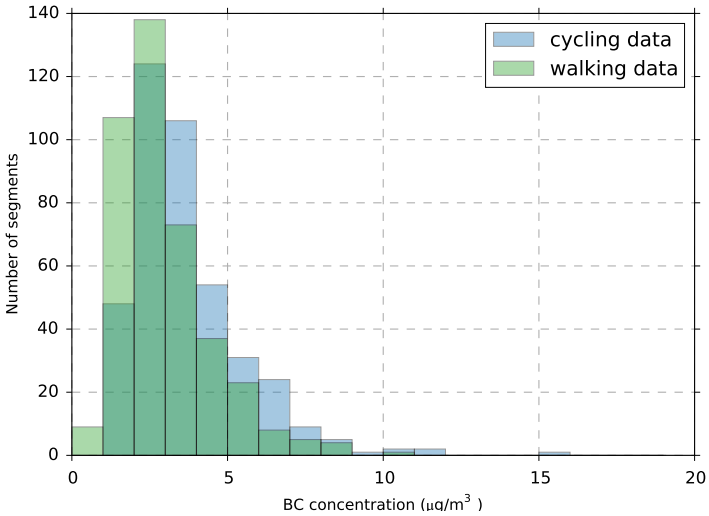


Figure 7.7: Histogram of the BC concentration levels (segment means) for cycling and walking data: two partly overlapping histograms indicate the shift towards higher values for the cycling data. Only those segments where data are available for both bike and foot modes are included.

cycling and walking data, but there are some differences in distribution over the day. The cycling subset has relatively more measurements in the morning (9-12 h), while the concentrations are typically higher in the morning than during the rest of the day (typically 18 % higher concentration in the morning compared to the afternoon). This can also contribute to the difference in concentration levels between cycling and walking data. The difference in distribution is, however, small (the morning hours constitute 53 % of the data for the cycling subset, and 45 % of the data for the walking subset) and does not fully explain the observed difference in concentration levels. Furthermore, when looking at the diurnal pattern, the difference in concentration levels is rather constant during the day. There is also no clear spatial pattern in the difference between cycling and walking data. As a last check, the difference is also consistently present when looking at each of the three teams separately.

We can conclude that there is an observable difference between the exposure during cycling and during walking, however, this difference is not large. For the following analyses in this chapter, we will combine the data of both modes of transport (unless explicitly stated).

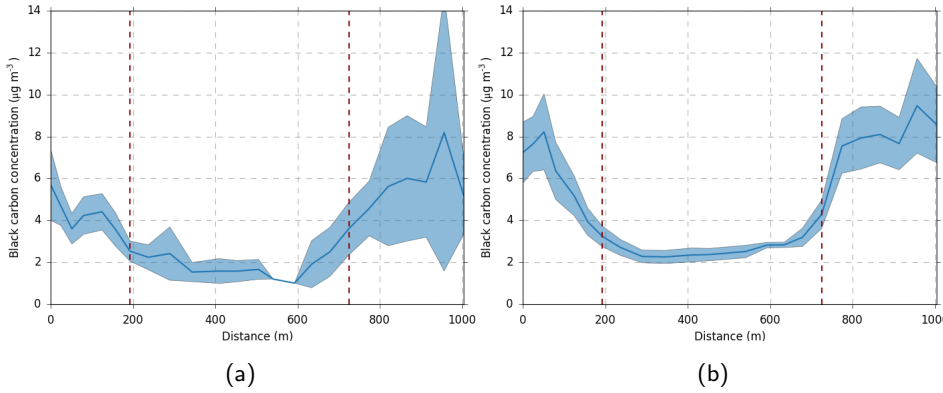


Figure 7.8: BC concentration profiles of the Stadspark and surrounding streets for (a) the campaign with city wardens and (b) the targeted campaign at 50 m resolution. Shaded area depicts twice the standard error on the mean. The vertical dashed lines indicate the borders of the park.

7.3.2 Comparison with targeted measurements

A previous case study in the same area was performed and described in the previous two chapters. For this study, an extensive data set of BC measurements was collected by bike in February-March 2012. This targeted campaign consisted of 256 and 96 repeated measurement tracks along two fixed routes that for a large part overlap with the study area of the opportunistic campaign described in this chapter (the route is indicated in Figure 7.2). Given the high number of repetitions, we can use this data set to benchmark the results obtained in this chapter.

In Figures 7.8 and 7.9, concentration profiles are shown for the Stadspark and surrounding streets and for the Plantin en Moretuslei. Broadly, the patterns in the concentration profiles correspond. For the Stadspark, the concentration increases sharply at the borders of the park towards the surrounding streets (with traffic) in both cases. For the Plantin en Moretuslei, the section between 200 and 600 m corresponds consistently to the section with elevated concentrations, although the exact locations of the highest values do not coincide in both campaigns. In the section between 600 and 1000 m, the Plantin en Moretuslei shows lower concentrations in both cases.

To quantify how good the results of both campaigns correspond, the overlapping 50 m segments are selected where at least 5 passages are recorded. For those segments the average concentration levels are calculated, and based on those values, the cor-

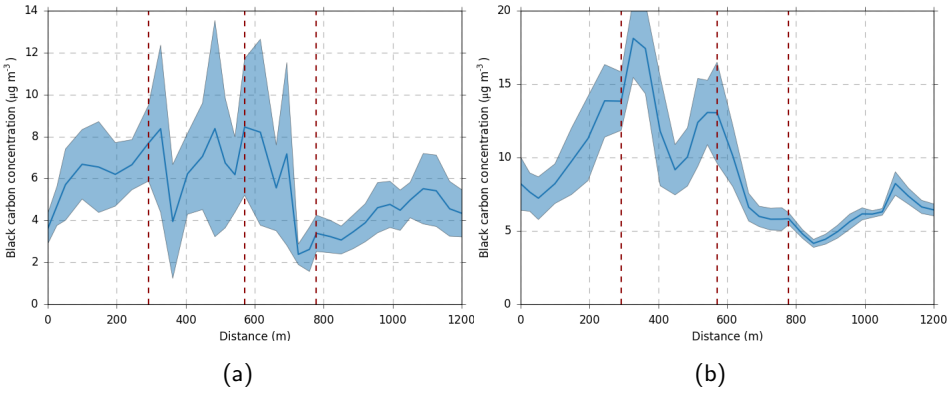


Figure 7.9: BC concentration profiles of the Plantin en Moretuslei for (a) the campaign with city wardens and (b) the targeted campaign at 50 m resolution. Shaded area depicts twice the standard error on the mean. The vertical dashed lines indicate the three major crossings along the Plantin en Moretuslei.

relation (Pearson's correlation coefficient), overall difference (difference in overall mean of all segments) and mean absolute deviation (MAD) are calculated.

First, a basic comparison without adjustments is performed, resulting in a correlation of 0.64, overall difference of 2.7 $\mu g\,m^{-3}$ and MAD of 2.8 $\mu g\,m^{-3}$. This comparison is also illustrated with a scatter plot in Figure 7.10. There is a large deviation in absolute value between both data sets: systematically lower values are found by the city wardens compared to the targeted campaign. Some possible explaining factors for this difference in concentration levels are a different mode of transport, period of the year, and period of the day during which the measurements are performed. Correction mechanisms for these factors are applied and a summary of the results of the comparison after correction is given in Table 7.1. Firstly, the targeted campaign only includes cycling data, while it was demonstrated in Section 7.3.1 that the bike data shows slightly higher concentration levels. When using only the cycling data for the current campaign, the overall difference decreases slightly to 2.5 $\mu g\,m^{-3}$. Secondly, the period of the day can also have an influence as the campaign with the city wardens also includes data in the afternoon, when the concentration levels are typically lower. Selection of the common hours (9-13h) for both datasets without performing a background normalization decreases the overall difference to 2.4 $\mu g\,m^{-3}$. Thirdly, the period of the year can have a large influence on the measured concentrations. This can be accounted for by using a background normalization to take the seasonal and day-

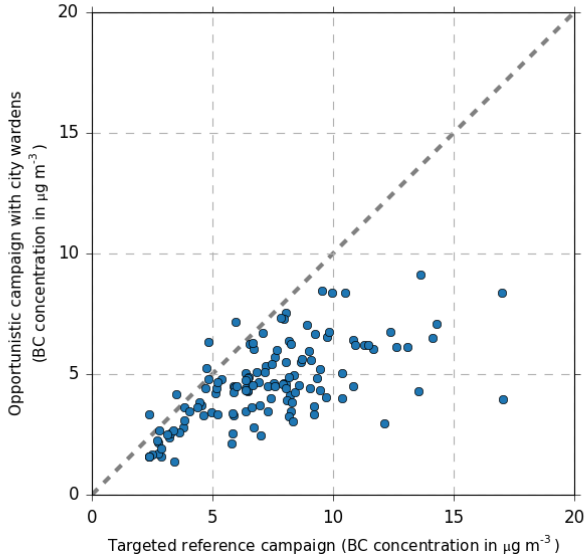


Figure 7.10: Comparison of the concentration levels in the overlapping 50 m segments between the campaign with city wardens and the targeted campaign

to-day variations into account. After performing a background normalization (see Section 7.2.4 on data analysis for more details), the overall difference decreases to 1.9 $\mu g\,m^{-3}$. However, when using a multiplicative background normalization instead of a combination, the decrease in overall difference is much stronger (down to a mean bias 0.6 $\mu g\,m^{-3}$). When combining these three correction mechanisms, the resulting overall difference is 1.7 $\mu g\,m^{-3}$ and MAD is 2.1 $\mu g\,m^{-3}$.

The same analysis is repeated for another spatial aggregation level, the street level. Instead of 50 m segments, the streets are used as aggregation level, which will result in segments of varying length. The result of the comparison at this level is also included in Table 7.1.

7.3.3 High spatial variability in air quality

The comparison in the previous section concerned only the part of the data set where the two campaigns overlapped. However, the spatial coverage of the campaign with city wardens was much larger. In this section, we show some results on the air quality for the entire study area.

Table 7.1: Comparison of the present opportunistic campaign with city wardens and the targeted campaign at two spatial aggregation levels (50 m segments and street level) and for different scenarios of data treatment. The evaluation statistics are the correlation (Pearson's correlation coefficient), overall difference (difference in overall mean of all segments or streets) and mean absolute deviation (MAD).

	50 m segments			Street level		
	Corr.	Overall diff.	MAD	Corr.	Overall diff.	MAD
Basic comparison	0.64	2.7	2.8	0.65	2.0	2.1
Only cycling subset	0.65	2.5	2.7	0.60	1.8	2.0
Common hours	0.65	2.4	2.5	0.53	1.7	2.0
Background normalized	0.64	1.9	2.2	0.65	1.2	1.6
Background normalized (mult.)	0.56	0.6	1.5	0.75	0.0	1.1
Combination of above	0.65	1.7	2.1	0.59	0.9	1.7

First, a map is constructed displaying the BC concentration for each 50 m segment in the study area where at least 5 passages have been recorded (Figure 7.11). The main roads clearly stand out, while the green areas and the more residential areas between main roads show much lower concentrations.

Going to a lower spatial resolution (all 50 m segments combined per street), Figure 7.12 shows a boxplot of the BC concentration for a selection of streets. Large differences between streets with high BC concentrations and streets with lower concentrations can also be noticed at this level (mean concentration levels ranging from around 1 $\mu g\,m^{-3}$ up to 5 to 6 $\mu g\,m^{-3}$). Typical streets with high concentration levels are busy roads with high traffic intensities (e.g. Turnhoutsebaan, Carnotstraat, Belgilei, Noordersingel) and street canyons (e.g. Provinciestraat). The lower concentrations can be observed at green areas (Stadspark, Krugerpark) and low-traffic residential streets (e.g. Van Leentstraat, Tuinbouwstraat).

7.4 Discussion

7.4.1 Sampling bias

The city wardens operated in a well-confined area in the city mostly during working hours on weekdays, resulting in repeated measurements. The spatial coverage within the study area is very high (Figure 7.3), although there is a large variation in the number of passages between different segments. For a lot of the streets and street segments only a small number of repeated measurements were recorded. The temporal coverage (Figure 7.4) also shows a non-uniformly spread pattern. Firstly, it shows that not all hours of the day and even not of the working day (9 am

Figure 7.11: Map of the BC concentration (50 m resolution): average concentration (trimmed mean) for each 50 m segment within the study area where at least 5 passages have been recorded.

- 17 pm) are measured proportionally. As a consequence of their work schedule, there are no measurements during rush hours when most people are commuting. Over the year, it can be noted that there are less measurements during winter. This can be caused by a lower motivation in the middle of the campaign, but also to worse weather conditions. Given the different typical concentrations during the day and the year, this will lead to sampling bias when targeting daily average concentrations.

Secondly, the temporal coverage also differs between locations: different locations are measured at different days and/or hours of the day. This leads to a different bias at varying locations, complicating the interpretation of the results. When the concentration levels at two locations are compared, a possible difference can also be partially allocated to sampling bias. To counteract this sampling bias by the time of the measurement, background normalization can be used. This will never fully cancel the differences introduced by sampling at other moments in time, but can improve the comparability. Background normalization was applied in the comparison to the targeted campaign, and is discussed there (Section 7.4.2).

The use of both walking and cycling modes of transport is another factor that complicates the interpretation of the air quality results. When dividing the dataset by mode of transport, there is a clear difference in average levels. After eliminating possible explaining factors (differences in location or time period of the measurements), the difference remains and it can be concluded that this difference

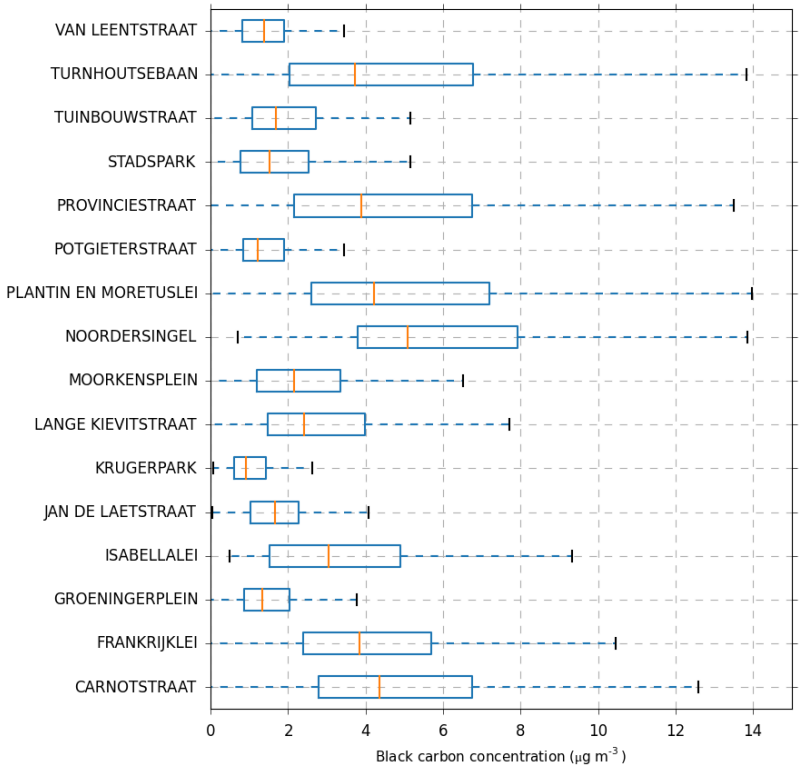


Figure 7.12: Boxplots of the BC concentration for a selection of streets (using all passages of the different segments in the specified street).

is systematic. This can be explained by the different distance to traffic (bike lanes are typically closer to traffic than the pavement) and similar results are reported in the literature (McNabola et al., 2008; Dons et al., 2012; de Nazelle et al., 2012). By combining both data sources, the mixed modes of transport will be an extra factor contributing to the variation in the dataset. Data of both modes combined represent the exposure during active mobility (soft modes of transport).

In general, the data from opportunistic measurement campaigns is biased and this consequently affects the resulting image of the air quality. It is not possible to exactly quantify the overall sampling bias, as the 'true' concentration levels are not known. However, correction mechanisms can be applied to reduce the bias.

7.4.2 Comparison of the unstructured opportunistic campaign and the targeted campaign

The resulting concentration levels of this unstructured opportunistic campaign were compared to those of the targeted campaign from the previous chapters. This comparison allowed us to investigate whether the approach with the city wardens could lead to comparable results and whether the sampling bias can be reduced or even eliminated. When looking at the concentration profiles (Figures 7.8 and 7.9) and given the correlation of 0.64 (Table 7.1), the larger spatial patterns (typical places with low or high concentrations, hotspots, ...) resulting from both campaigns correspond.

But despite this correspondence, both a large overall difference and a large variability at the individual streets and street segments is present. The large overall difference in concentration levels is caused by the different factors contributing to the sampling bias, i.e. different period of the year, different period of the day and different mode of transport. Of course, other factors could also contribute to the sampling bias (e.g. instrumental bias), but these are not further addressed in this study. Using only the cycling data slightly improves the comparison (reduction of the overall difference), but also decreases the number of measurements that can be used in the comparison, making the comparison less reliable. The same can be said for selecting the common hours.

The background normalization is used to counteract the bias introduced by measuring on different days, and when using hourly reference values, this can also tackle the bias introduced by measuring at different hours of the day. Applying the background normalization has a higher positive impact on the overall difference, but this depends largely on the used algorithm. In literature, background

normalization is a commonly used technique for temporal adjustment of measurements that are not taken simultaneously (e.g. Hoek et al., 2002; Zuurbier et al., 2010; Eeftens et al., 2012b; Tan et al., 2014). The adjustment can be additive (difference) or multiplicative (ratio), and it is generally stated that the two methods do not give very different results (Hoek et al., 2002). In the cited studies, the additive method was used. Dons et al. (2012) proposed a combined method to prevent negative values. These findings contrast with the result seen here, were the multiplicative method performs much better than the combined method. The multiplicative method is more suited to correct for temporal variation driven by meteorological conditions, while the additive method corrects for the influence of regional pollution (Tan et al., 2014). BC is mainly coming from local sources (traffic) in the study area, which is a possible reason that the multiplicative method works better in this study. It is clear that, although the correction mechanisms could reduce the difference between both data sets, a considerable difference between the targeted and opportunistic data set remains. Using a sampling experiment, Tan et al. (2014) also showed that correcting measurements from short duration campaigns with data from continuous stationary monitoring sites does not completely remove the uncertainty due to short duration sampling.

Besides the overall difference, there is a large variability on the individual segments (expressed by the mean absolute deviation, MAD). The artefacts are due to the short time period of the individual mobile measurements and the highly variable air quality concentrations at short temporal resolution, combined with a limited number of repeated measurements. The city wardens campaign still has not sufficient repetitions for all segments, and so a higher variability on the individual segments is to be expected. Most segments are measured between 9 and 27 times, while in the previous study on the targeted campaign it was concluded that 24 to 94 repetitions are required to map the BC concentrations at a 50 m resolution with an uncertainty of 25 % (5 to 11 repetitions were required when relaxing this constraint to 50 %, see the previous chapter). We can conclude that not enough data were gathered to assess the air quality at such a high spatial resolution. When looking at a lower resolution (street level), the agreement between both data sets is better.

Despite the systematic difference between both data sets, we can conclude that the larger spatial patterns are captured by the opportunistic campaign, but that there still remains a large uncertainty on the individual segments.

7.4.3 Evaluation of the monitoring campaign with city wardens

Thanks to the collaboration with the city of Antwerp, we had the opportunity of setting up this campaign with the city wardens. The involvement of the environmental agency of the city and their interest in the results made it easier to engage city personnel for performing the measurements. Working with the city wardens offered some benefits compared to volunteers or other professionals. Because they are city personnel, all instruments were always connected and recharged at their office and the follow-up could be centralized. Further, the city wardens went to varying places in the city as indicated by the high spatial coverage (Figure 7.3), e.g. also visited traffic-free and green areas. Other examples of carriers such as postmen or parking wardens can partly have the same advantages (involved through their job), but can have a more limited temporal or spatial coverage. For example, postmen or parking wardens will typically less likely visit traffic-free or green areas.

Some drawbacks or aspects that could be improved in the monitoring approach could be identified. A lot of variability was introduced due to several factors such as varying modes of transport, sampling during different hours of the day and of the year (see Section 7.3.1). This allowed us to evaluate the potential of a case study that manifests such characteristics, but for the sake of this campaign itself some optimizations could have been performed. A stronger follow-up of the wardens could have kept the level of measurements more constant during the year. The mixed mode of transport of the wardens is an additional complication and it should be avoided if possible (which was not the case in this campaign).

Another problem is the low GPS quality in urban environments, and consequently an intensive processing of these data is needed. In contrast to the targeted monitoring case, no a priori knowledge on the tracks that are monitored is available. Combined with the low GPS quality, this complicated the detailed mapping of the air quality. Due to a possibly large deviation from the true position, some measurements may not have been assigned to the correct street. Also, data are not only collected outdoors. It turned out to be hard to detect, in a reliable way, when a city warden enters a building. We have tried to tackle this by filtering the data on the number of satellites, but this may have led to some excessive data cleaning of good measurements on the one hand or the incorrect use of indoor measurements on the other hand. Possible solutions are the use of better GPS devices (but this can become expensive), better map-matching or transport mode detection algorithms or the use of diaries (this could give more contextual infor-

mation to interpret the GPS data, but it also puts a much higher burden on the people carrying out the measurements).

A next issue is the instrument used to measure the air quality. The micro-aethalometer signals may sometimes be noisy or show instabilities. This will be an extra factor contributing to the uncertainty on the results. Also, the currently used bag does not protect the micro-aethalometer from the rain. A better system would make it possible to also measure when it is raining.

Collecting data consistently during a one year period by the city wardens is quite a challenge and turned out to be less evident than expected. Good usability of the monitoring equipment is crucial. The measurement set-up needs to be designed to keep the impact on the city wardens limited. Tasks that look simple at first such as making sure the battery of the measurement device is recharged, changing a filter, or turning off measurement devices after completion of the measurement day seem to take more effort than expected. The city wardens seemed to conceive the monitoring as an extra task to their daily job. A stronger follow-up could have kept the data collection at a more constant level throughout the year.

7.4.4 Relevance for participatory sensing and air quality applications

With the set-up of this campaign, we somehow mimicked an uncoordinated participatory monitoring campaign with actually available sensors. In the future, mobile participatory campaigns could potentially manifest much of the same characteristics of this data set in terms of uneven spatial and temporal coverage and different modes of transport with increasing numbers of data as monitoring equipment becomes cheaper. In this way, this case study emulates the kind of data that would be gathered from participatory campaigns in an achievable way for the time being. We demonstrated that such an approach can indeed be used to identify broad spatial trends and hotspots over a wider area. The results show that the spatial patterns in the urban BC concentrations can be assessed, obtaining a spatial resolution which would not be possible with stationary monitoring stations. This shows that the approach can be used for air quality applications including hotspot identification, personal exposure studies, regression mapping (using the results to build land-use regression models), etc. However, it is clear that repeated measurements and careful processing and interpretation of the results are very important. Even with a rather large number of measurements (393 h), there is still a large variability on these results. This indicates that an important challenge for

unstructured participatory monitoring campaigns will be to collect sufficient data to cover both spatial and temporal variability.

7.5 Conclusion

We defined *opportunistic mobile monitoring* as data collection making use of existing carriers to move measurement devices around. Based on an opportunistic mobile monitoring campaign with the collaboration of city employees (city wardens) carrying micro-aethalometers during their daily surveillance tours, we were able to map the exposure to black carbon in the urban environment. A high spatial coverage in the selected area was obtained during the one-year data collection. Due to the nature of the wardens' work, there is no control over the route followed. This results in an unstructured data collection both in space and time, leading to sampling bias. A temporal adjustment can partly counteract this bias, but will not fully remove the uncertainty associated with the temporal variation. There is still a rather large uncertainty on the average concentration levels at the individual locations (when looking at a high spatial resolution of 50 m) due to a limited number of measurements and sampling bias. Despite this uncertainty, large spatial patterns within the city are clearly captured. In anticipation of lower-cost sensors, this case study with city wardens demonstrates a promising approach to monitor urban air quality at a high spatial resolution based on opportunistic data collection.

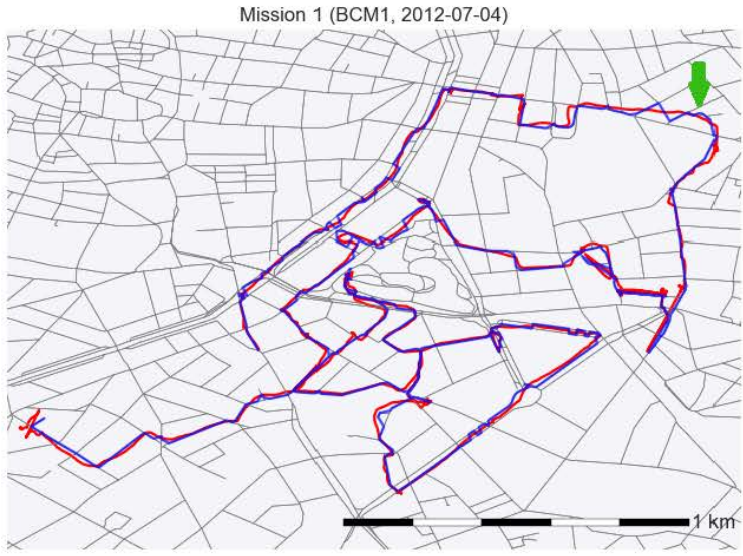
This chapter illustrates the potential and limitations of participatory monitoring approaches. With the set-up of this campaign, we somehow mimicked an uncoordinated participatory monitoring campaign with the currently available means. If appropriate low cost instruments would be available, citizens can carry it around measuring the air quality in a non-coordinated way. The results in this chapter demonstrate that such an approach can indeed be used to identify broad spatial trends over a wider area, enabling applications including hotspot identification, personal exposure studies, regression mapping, etc. But they also emphasize the challenges with unstructured opportunistic mobile monitoring. It is clear that repeated measurements and careful processing and interpretation of the results are very important. Unstructured participatory monitoring campaigns will have to collect sufficient data to cover both spatial and temporal variability. Until low-cost air quality sensors are available that can collect data of sufficient quality in an effortless way, continued motivation and follow-up of participants is crucial.

7.A Supplementary data

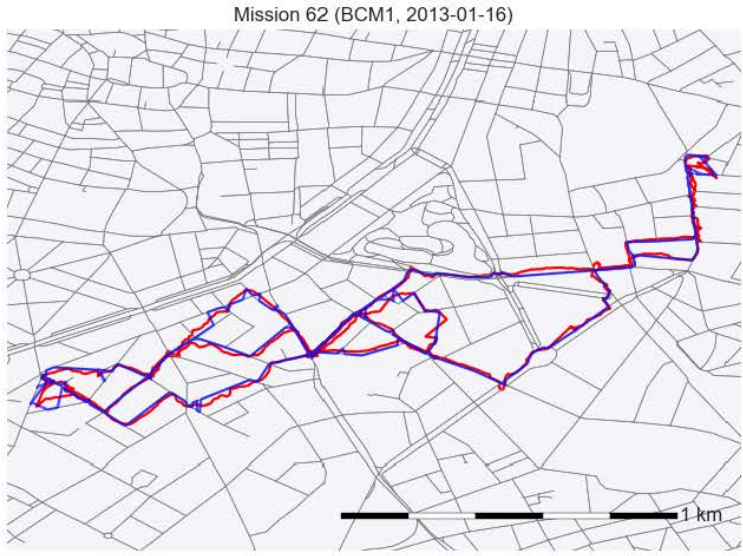
7.A.1 Example GPS tracks

Two example GPS tracks are included as an illustration of the typical GPS tracks gathered by the city wardens (Figure 7.S1). The tracks show that in general the map matching procedure finds the correct streets. However, sometimes errors will occur. For example, in the cycling track (Figure 7.S1a), the city warden probably followed the Carnotstraat, although the mapped track follows the neighbouring streets (see the green arrow indicated on the figure). The occurrence of such errors is unavoidable, although a more advanced map matching algorithm may reduce the frequency of occurrence.

Based on a visual analysis, the GPS track of the city warden on foot seems a bit more shaky compared to the cycling track, although the quality of the map matching will not likely be worse. It is difficult to assess the quality of the GPS tracks based on the distance between the original track and the map matched track, as the expected distance between both tracks (apart from the GPS error) will vary from location to location and between transport mode (e.g. the bike lane is typically closer to the road, and hence closer to the street from the street database, as compared to the pedestrian area).



(a)



(b)

Figure 7.S1: Example GPS tracks for a day with predominantly a) cycling and b) walking measurements. The red line is the raw track, the blue line map matched track.

CHAPTER 8

Air quality mapping based on mobile monitoring: guidelines and applications

8.1 Introduction

In the previous chapters, the use of mobile monitoring to map the local urban air quality at high resolution has been evaluated. Here, local air quality refers to the average outdoor air pollution people are exposed to in urban environments. It is not limited to urban baseline concentrations, but also includes very localized effects, which are measured in this thesis at a resolution of 20 to 50 m (below 100 m). This high-resolution mapping can be used to detect local hotspots, to calculate activity-based exposure to air pollution, or to validate high-resolution models, amongst other applications.

In this chapter, the experience from the previous chapters is collected in practical considerations and guidelines for mobile monitoring campaigns to map urban air quality. Further, some examples of mobile monitoring that are related to this thesis are presented: a case study in Oslo and applications of the airQmap platform.

8.2 Practical considerations and guidelines for mapping urban air quality using mobile monitoring

Mobile monitoring is often presented as an additional monitoring tool to increase the spatial resolution of air quality maps. However, an insufficient number of mobile measurements may only represent snapshots of the air quality and is not necessarily representative. Therefore a careful set-up of the mobile monitoring campaign is needed, and processing and interpretation of the mobile data deserve increased attention.

- The occurrence of events can lead to considerable bias when using the arithmetic mean to aggregate mobile data, certainly when dealing with a smaller sample size. Appropriate actions to investigate this bias are to assess the possible impact of events on the results in relation to the goal of the campaign and to investigate the results from different statistics (arithmetic mean, median, trimmed mean).
- The trimmed mean resulting from the removal of the lowest and highest measurement at the spatial aggregation level (i.e. for each segment) is proposed as an integrated outlier detection method to reduce the detrimental impact of random events on the calculated averages. This approach does not over-correct for systematically occurring peaks that can be typical at certain locations or streets and is a better estimator for the average concentration (i.e. the average including the tail measurements) than the arithmetic mean when dealing with a limited number of repeated measurements.
- To account for the temporal variation in the background concentrations, a temporal adjustment based on these background concentrations can be applied. The temporal adjustment reduces the variability between different runs and leads to more reliable results with lower monitoring efforts. Especially for campaigns with an unstructured set-up where not all locations are measured at approximately the same moment, the temporal adjustment can partly counteract the sampling bias.
- Spatial aggregation is applied to smooth the data at different spatial levels (routes, streets, segments). We showed that mapping air quality at a spatial resolution up to 50 m is feasible for BC. A higher spatial resolution of 20 m can be obtained with a slightly increased uncertainty depending on the sample size.

- To generate high-resolution maps over larger areas (several km^2), large quantities of data are required to include the range of possible meteorological conditions and the range of local air quality conditions (depending on local sources, e.g. traffic intensity), and to counter occasional and exceptional events. It is important to assess whether enough repetitions are made in relation to the goals of the monitoring campaign.
- The data requirements can be reached by performing repeated monitoring runs. The exact requirements depend on the spatial and temporal resolution of the map and the quality constraints of the results, but can be decreased by using a trimmed mean for averaging and by performing a temporal adjustment. In the targeted case study, 10 repeated measurement runs will be sufficient to estimate the average concentration at almost all locations when targeting an uncertainty of 50 %. When performing 40 repeated measurement runs, the average concentration could be estimated at half of the locations with an uncertainty of 25 % or below. For an unstructured opportunistic campaign, the numbers will probably be higher due to the additional uncertainty caused by sampling bias.

The above guidelines are generally valid for both targeted and opportunistic measurement campaigns. However, there are some differences between both approaches. In case the opportunistic campaign has a structured set-up (following fixed routes), the characteristics of the obtained dataset will be similar to a dataset from a targeted campaign. Regarding the processing and interpretation of the data, the main difference is between structured and unstructured data collection.

In a campaign with a structured data collection, all measurement locations are monitored at approximately the same time (or within a short time period, i.e. the time it takes to complete one measurement run). This ensures that the temporal variation (day-to-day variation, rush hour versus non-rush hour) has less influence on the differences in measured concentrations between locations. In this case, a temporal adjustment is not essential to counter the sampling bias caused by the time of the measurement. A temporal adjustment can still be used to extrapolate the measured concentrations to e.g. a typical yearly average. This extrapolation introduces additional uncertainty on the obtained value, but the adjustment is needed when one wants to compare the concentration levels to other campaigns that are conducted in other periods. In case of an unstructured opportunistic data collection, different locations will possibly be measured at different days and/or hours of the day, which introduces sampling bias. Temporal adjustment can then be used to reduce this sampling bias. However, this adjustment is not perfect and

unstructured data collection will therefore typically result in data with a larger uncertainty compared to structured data collection.

Limitations

There are also some important limitations to the mobile monitoring approach. First, mobile data are representative for the time period in which they were collected, which is typically a confined time frame. A targeted campaign is typically deliberately planned at certain moments in a limited time period. For example, the targeted campaign in this thesis was conducted each day between 7 am and 13 pm for several days within a period of 4 weeks. In an opportunistic campaign, it can be easier to maintain the measurements for longer time periods and obtain a larger temporal coverage. However, also opportunistic data collection is often limited to, for example, rush hours (e.g. commuting routine) or daytime moments (e.g. tram schedule). As mentioned above, a possibility is to use temporal adjustment to extrapolate the results to other periods assuming that the local contributions are similar in time and that the main temporal variation is caused by background effects that are normalized.

Second, the representativeness of the result depends on the type of mobile platform. In the targeted case study in this thesis, a bicycle was used as monitoring platform and hence the results are representative for the exposure of cyclists, but not necessarily for other transport modes such as pedestrians and car drivers, or for residents.

Third, mobile measurements are typically limited to measurements along the street. Air pollution levels can, however, also vary considerably orthogonally to the street direction. The type of mobile measurements used in this thesis will not capture this variation, and it can even be the cause of part of the observed differences between locations (e.g. varying distance of the biking lane to traffic, different predominant transport mode at different locations).

Fourth, measurements of BC are used in this study. Pollutants such as BC, UFP and NOx are strongly traffic-related and will therefore show a high variability in urban environments. Other types of pollutants such as $PM_{2.5}$ or PM_{10} have lower small-scale dynamics and the guidelines will differ for these other pollutants.

Fifth, despite the strategies to reduce the required number of repetitions, a large number of repeated measurements is still required. A targeted case study requires dedicated man-power to carry out the measurements, which can be very time consuming. Such case studies can also be conducted with the involvement of a

limited number of volunteers, but this requires highly committed participants. Opportunistic approaches can reduce this workload taking advantage of existing infrastructure or using the normal daily routines of people.

Last, a monitoring approach will always be limited in space and will therefore not result in a map with full spatial coverage. Especially a targeted campaign focuses typically on a limited number of streets. Depending on the specific set-up, an opportunistic campaign can result in a higher spatial coverage. When lower-cost sensors with sufficient performance become available, data can be collected on a larger scale in an uncoordinated, opportunistic way. This way, participatory monitoring has the potential to deliver the high amounts of data needed for obtaining representative maps with high spatial coverage based on mobile monitoring.

8.3 airQmap – DIY air quality mapping

airQmap[1] is a platform developed by VITO that is based on the methodology developed in this thesis. It is an application of a targeted mobile monitoring strategy, based on repeated measurements on fixed routes spread over different days (20 to 30 repeated measurement runs are typically performed). The platform allows people with limited or no air pollution expertise, such as city personnel or volunteers, to carry out air quality measurements and to get a detailed view on the air quality at the street level. VITO offers airQmap as a commercial service to interested parties (e.g. cities and municipalities).

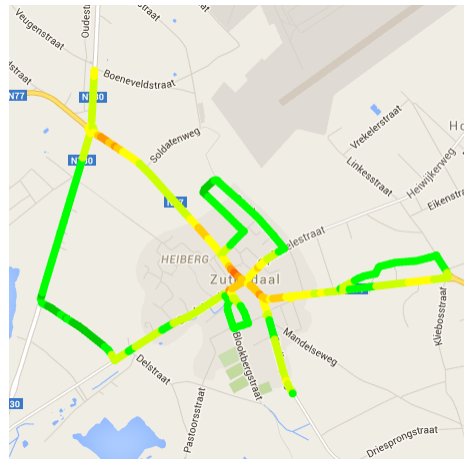
The airQmap platform contains two parts. The first part consists of easy to use measurement devices: the micro-aethalometer to measure BC, a GPS and a netbook to synchronize the instruments and transmit the data. The second part is an automated data processing infrastructure which constructs a BC map from the transmitted data. The processing infrastructure is an implementation of the guidelines and includes quality control and processing of the BC and GPS data and spatio-temporal data aggregation. Further, airQmap provides an interactive map that can be consulted through the web.

The airQmap platform has been successfully used in cities such as Antwerp, Ghent, Brussels, Liege, Amsterdam and Oslo, and in Belgian municipalities such as Mol, Beringen and Zutendaal. Some of the resulting maps are shown in Figure 8.1 and 8.2. In Oslo, a measurement campaign was set up with the collaboration of the Norwegian Institute for Air Research (NILU). The measured BC concentration

[1] `http://www.airqmap.com`

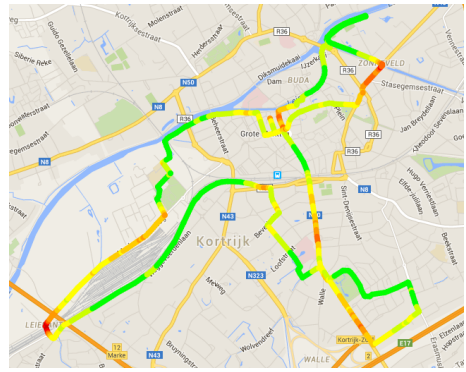
Figure 8.1: Map of the BC concentrations based on mobile measurements in the area of Majorstuen, Oslo (aggregated to a spatial resolution of 50 m).

(a) Beringen



(b) Zutendaal

(c) Amsterdam



(d) Kortrijk

Figure 8.2: Black carbon maps for some cities and municipalities using the airQmap platform (from: http://airqmap.com/).

levels are shown in Figure 8.1 and the concentration map indicates a strong spatial variability. The campaigns in Antwerp, Ghent and Brussels were conducted by Test Aankoop. In Amsterdam, the campaign was a collaboration between VITO and the Dutch Longfonds, the European Lung Foundation and the European Respiratory Society as an initiative for the ERS/ELF Healthy Lungs for Life campaign (Figure 8.2c) (Theunis et al., 2015). The campaigns in Beringen and Zutendaal were conducted by the municipality with the help of volunteers. The results of the different measurement campaigns confirm that mobile monitoring is an interesting method to characterize the spatial variation in air pollutant concentration levels.

8.4 Conclusion

Given a careful set-up and data processing, mobile monitoring campaigns are a valuable tool to get a detailed view on the air quality and the exposure of citizens at street level. The multiple measurement campaigns presented in this chapter and the interest of local governments and other organisations show the added value of this measurement-based approach. Targeted monitoring campaigns can be used to conduct a sufficient number of repeated measurement runs in a well-defined area to satisfy the data requirements. Alternatively, opportunistic campaigns can be set up that require less dedicated man-power to carry out the measurements. But, with unstructured opportunistic data collection it can be more difficult to obtain sufficient repeated measurements at each location, and sampling bias leads to additional uncertainty.

A measurement-based approach is limited and it is typically not possible to obtain a pollution map with full spatial coverage. Targeted campaigns, such as the examples of airQmap, are certainly not covering a full area, but also with opportunistic data collection it remains difficult to get a full spatial and temporal coverage. Interpolation methods or regression models can be used to increase the coverage of the pollution map, and this will be investigated in the next part of this thesis.

PART III

SPATIAL AND TEMPORAL EXTRAPOLATION

CHAPTER 9

Land use regression models for drawing air pollution maps

9.1 Introduction

City-wide maps of the outdoor air pollution are typically used in epidemiological studies as proxy for personal exposure in order to overcome the limitations of fixed-site monitoring and personal monitoring. But up-to-date information on urban air quality with high spatial and temporal resolution is also important to inform air-quality policy and traffic management, and to inform the general public (Briggs et al., 2000; Jerrett et al., 2005; Beelen et al., 2009; Vardoulakis et al., 2011; Beelen et al., 2013).

In the previous chapters, methods have been developed to monitor air pollution with a high spatial resolution based on mobile measurements. However, these methods do not result in city-wide pollution maps with a full spatial and temporal coverage. Widespread deployment of the relatively high quality sensors such as the micro-aethalometer is still hampered by their high purchase cost and, in some cases, the manpower involved to perform the measurements. But also using low-cost sensors relying on a crowd-sourcing strategy does not automatically lead to complete data sets with uniform coverage.

A range of methods exists to go beyond the spatial and temporal coverage of the measurements and draw pollution maps. These methods include empirical methods based on measurement data, such as interpolation or land use regression, and deterministic models, such as dispersion models. This chapter presents a brief

overview of these different methods, with special focus on land use regression as it will be used in the next two chapters.

9.2 Methods to draw air pollution maps

Interpolation Actual pollution measurements can be interpolated using several methods including inverse-distance weighting or geostatistical methods. The most common geostatistical method used in the air pollution field is kriging (Jerrett et al., 2005). However, these methods often do not work very well. Especially in urban environments, the pollution levels show very complex patters, due to traffic dynamics and street configuration, and with steep gradients and localized hotspots. To describe this adequately, interpolation methods require a dense network of measurement sites (Briggs, 2005; Beelen et al., 2013; Can et al., 2014). It is possible to include additional variables such as land use variables in the interpolation process (e.g. universal Kriging, co-Kriging) to improve the predictions. However, according to Briggs (2005), it is often found that the additional variables seem to account for most of the explicable spatial variation in concentrations, and equally good (or better) results can be achieved using these variables in multivariate regression models.

The RIO model is an example of a more advanced interpolation model, and is used in Belgium to interpolate the measurements from the fixed monitoring stations. It can be classified as a detrended Kriging model (Janssen et al., 2008). In a first step, the local character of the air pollution measurements is removed in a detrending procedure. Subsequently, the detrended data is interpolated using ordinary Kriging to a country-wide 4 by 4 km^2 grid. Finally, in a re-trending step, a local bias is added to the interpolation results based on land use (CORINE) variables. An example output, the annual mean NO_2 concentrations of 2014, is shown in Figure 9.1. Given the coarse grid, this model cannot be used to characterize small-scale within-city variability.

Land use regression (LUR) LUR modelling tries to relate the pollution concentration at a given site with surrounding land use and traffic characteristics using a regression model, typically a multivariate linear model. The main strength of LUR is the empirical structure and its relatively low cost, but it is limited when moving to study areas with different land use (Jerrett et al., 2005). This method is discussed in detail in the next Section 9.3.

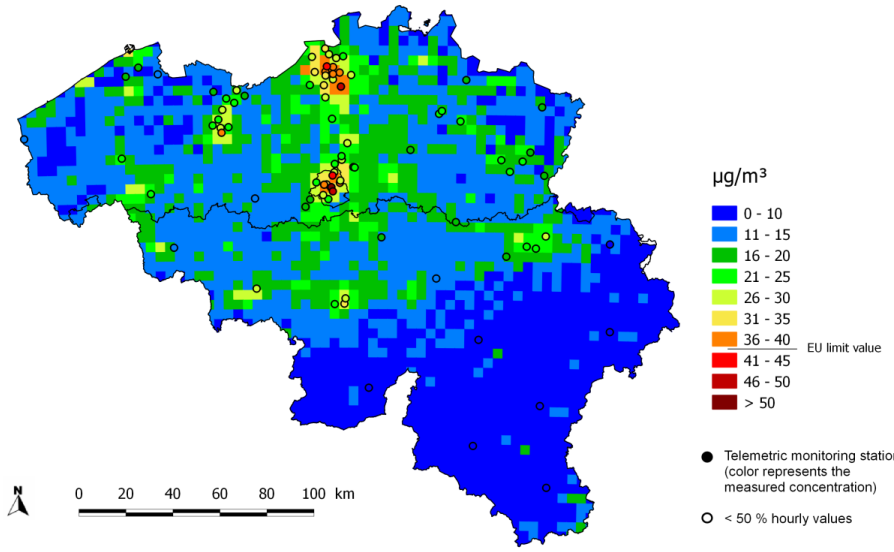


Figure 9.1: Annual mean NO_2 concentrations (Belgium, 2014) based on the RIO model (IRCEL, 2015).

Dispersion models In contrast to the previous methods, dispersion models do not start from air pollution measurements. Dispersion models are based on assumptions about deterministic processes making use of data on emissions, meteorological conditions, and topography in estimating spatial air pollution concentrations. Dispersion models have the potential advantage of incorporating both spatial and temporal variation of air pollution without the need for dense monitoring networks. On the other hand, these models need relatively costly and detailed input data, require a relatively high level of programming expertise coupled with expensive hardware requirements, and make a lot of assumptions about dispersion patterns (Jerrett et al., 2005; Beelen et al., 2009, 2013). Further, dispersion models typically do not result in street-level concentrations, unless the model is explicitly coupled with a street canyon model (Lefebvre et al., 2013).

In addition, Jerrett et al. (2005) also distinguish hybrid models. These models combine multiple methods, including personal or regional monitoring. For the global burden of disease analysis, Brauer et al. (2016) used a combination of satellite-based estimates, chemical transport model simulations and ground measurements to produce global estimates of annual average $PM_{2.5}$ concentrations. Lefebvre et al. (2013) presented another example of a hybrid model, the RIO-IFDM model chain. It consists of the interpolation model RIO mentioned above, coupled with a dispersion model (IFDM, a bi-gaussian plume model). It can be

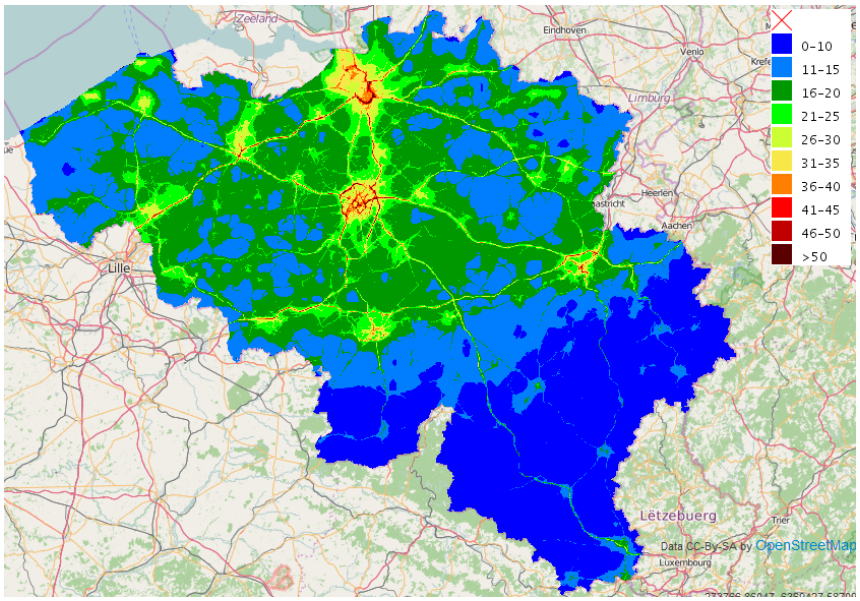


Figure 9.2: Annual mean NO_2 concentrations (Belgium, 2013) based on the RIO-IFDM model (source: http://www.atmosys.eu).

further combined with a street canyon model to simulate street-level concentrations (Lefebvre et al., 2013). An NO_2 map generated by this model is shown in Figure 9.2.

Several studies compare interpolation, LUR and dispersion modelling for exposure assessment in epidemiological studies. However, contrasting results can be found in literature. In general, it has been shown that LUR outperforms geostatistical methods (Hoek et al., 2008). Comparisons between LUR and dispersion models often conclude that both yield similar results (Cyrys et al., 2005; Wang et al., 2015). de Hoogh et al. (2014) found a high correlation between a LUR and a dispersion model when estimating small-scale variations in NO_2 concentrations. In a study of Beelen et al. (2010), on the other hand, a dispersion model explained concentrations at validation sites better than a LUR model. Compared to dispersion models, LUR modelling requires less detailed input data at the expense of the need to obtain monitoring data for a sufficiently large number of sites (Hoek et al., 2008). When using LUR, the model learns about the study area and the sources based on the measurements, whereas dispersion models take this information from other sources (e.g. emission data for different pollution sources).

9.3 Land use regression models

Land use regression tries to relate the air pollution concentration level at specific locations with predictor variables, usually obtained through geographic information systems (GIS) and holding information on surrounding land use and traffic characteristics, using a regression model (Jerrett et al., 2005; Hoek et al., 2008) This model can then be used to predict the pollutant concentration at unsampled sites. The technique was initially called regression mapping and introduced in the SAVIAH (Small Area Variations In Air quality and Health) study in the late 90s (Briggs et al., 1997). LUR models are widely used to map the spatial variability in air pollutant concentrations, especially for small-scale within-city variation (Jerrett et al., 2005; Hoek et al., 2008; Beelen et al., 2013). They are used to estimate individual air pollution exposure in epidemiological and risk assessment studies (Hoek et al., 2008; de Hoogh et al., 2014).

The development of the models is based on measurements of air pollution at typically 20–100 locations spread over the study area. Measurements from the routine network can be used, but in most cases measurement campaigns are specifically undertaken for the purpose of LUR modelling. LUR methods have been applied to model annual mean concentrations of different pollutants including NO_2, $PM_{2.5}$ and BC, and in different settings, including European and North-American cities (Hoek et al., 2008). Due to the availability of inexpensive passive samplers for NO_2, the majority of LUR models were initially being developed using NO_2 measurements as an indicator of traffic-related air pollution (Hoek et al., 2008; Larson et al., 2009). In recent years, LUR models have also been developed for particulate matter. For example, within the ESCAPE (European Study of Cohorts for Air Pollution Effects) project, particulate matter is monitored using active samplers next to passive NO_2 samplers to develop LUR models (Eeftens et al., 2012b; Wang et al., 2013). Dons et al. (2014a) also monitored BC with an active sampler in addition to NO_2.

9.3.1 Model development

Typical predictor variables that are used in the development of LUR models include: traffic variables (traffic intensity, road length, distance to roads), land use variables (residential, industrial, port, airport, green urban areas), population density, topography (e.g. altitude), physical geography (e.g. street canyon indicators), meteorology and location (coordinates). The predictor variables can be calculated

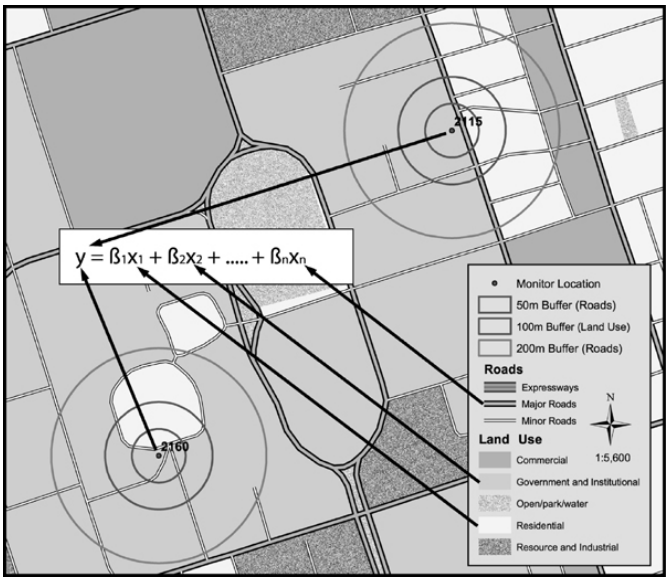


Figure 9.3: The concept of a land use regression model showing sampling locations as response variable and land use characteristics within buffers as predictor variables (Jerrett et al., 2005)

as a point estimate or based on a certain buffer around the measurement location (Figure 9.3).

Most LUR studies use standard multiple linear regression techniques to develop prediction models (Hoek et al., 2008). Variable selection methods are used to build a model from the large set of available predictor variables. A supervised stepwise approach as described in Henderson et al. (2007) is widely used (Hoek et al., 2008; Eeftens et al., 2012a). This approach focuses on selecting models with a limited number of predictor variables that have plausible (predefined) effects. It is a stepwise forward search for the best subset of predictor variables, based on an optimization of the adjusted R^2 and on the significance of the coefficients in the linear regression model. Often there is a supervised step included as a check of the direction of the effect. The motivation of this supervised approach is to focus on plausibility and interpretability of the model and to limit the risk of overfitting. In this way, the model could also be applied more easily to other study areas (Hoek et al., 2008; Beelen et al., 2013).

Other techniques for selecting the predictor variables exist, including forward or backward searches. In the above described approach (supervised stepwise), the variable selection in the model-building process is based on all data. Alternatively, the variable selection can also be based on the performance during cross-validation,

using e.g. the cross-validation R^2 instead of the adjusted model R^2, to select those variables that give the best generalization and to minimize the risk of overfitting. For example, Su et al. (2015a) applied a deletion/substitution/addition (DSA) algorithm that searches through the variable space in order to minimize the squared prediction error during cross-validation. This DSA algorithm removes much of the judgement of the researcher by only using model cross-validation as a model evaluation measure leaving the variable selection to the DSA algorithm (Beckerman et al., 2013). Basagaña et al. (2012) compared different techniques (the supervised approach with a forward selection based on adjusted R^2, the same algorithm but forward selection based on cross-validation R^2 and the DSA algorithm). They concluded that the algorithms performed similarly in terms of predictive ability on the validation dataset.

Most studies make use of linear models. Some relations between the response (pollutant concentration) and the predictor variables are, however, not linear. This can be addressed by introducing non-linearities manually, e.g. by adding the inverse distance to a road as a predictor variable. Some studies use non-linear techniques such as Generalized Additive Models (GAM) (e.g. Li et al., 2012; Hasenfratz et al., 2015; Dekoninck et al., 2015).

9.3.2 Model validation

Evaluation is an essential part of model development. To assess how well the model predicts at unsampled locations, the model R^2 is too optimistic since the model is optimized for the used data, especially for models based on small datasets and a large number of possible predictor variables. In literature, LUR models often use leave-one-out cross-valiation (LOOCV) to assess the model performance (e.g. Hoek et al., 2008; Eeftens et al., 2012a; Beelen et al., 2013), but it is known that this may also overestimate the predictive ability on independent data sets (Wang et al., 2013). Some other studies used hold-out validation (HV) or random n-fold cross-validation to get a more reliable estimate of the performance. In the hold-out set-up, part of the data is excluded for model building, and the performance of the fitted model is tested on this independent validation data set. This procedure can also be applied in a cross-validation scheme by repeatedly selecting a random subset of the data as validation data (e.g. Hasenfratz et al., 2015; Kanaroglou et al., 2013; Montagne et al., 2015).

The overestimation of the model performance of a LOOCV R^2 is especially apparent for LUR models based on a small number of sampling locations. Recent studies show that such models can perform well in the LOOCV setting, but per-

form poorly in evaluation on an independent dataset (Wang et al., 2012; Basagaña et al., 2012; Wang et al., 2013). Wang et al. (2013) investigated the impact of the cross-validation methodology on the reported performance for different study areas. For a model based on 20 sites predicting $PM_{2.5}$ absorbance, the median LOOCV R^2 was 0.81, whereas the median HV R^2 was 0.44. Basagaña et al. (2012) investigated the impact of the sample size on model performance. They calculated the model R^2, LOOCV R^2 and HV R^2 for different sample sizes (using a sampling exercise from a larger dataset). With a sample size of 20, they even found negative values for R^2 (explained variance of around 0.20) for the validation dataset (HV), while the model R^2 values were around 0.70–0.80. For larger sample sizes (up to 120 samples), the difference became smaller: LOOCV R^2 decreased and HV R^2 increased with increasing number of sampling sites, but the LOOCV R^2 remained an overestimation. The higher HV R^2 indicated that the models developed using more sampling sites were more robust, despite their lower model R^2. Basagaña et al. (2012) concluded that LUR models aiming at characterizing local air pollution levels in a complex urban environment should be based on a large number of measurement sites (> 80 in their setting).

Caution should be taken when interpreting and comparing R^2 values, given the impact of the evaluation method as explained above and the different definitions of the metrics used throughout literature. Furthermore, often all data are used in the variable selection process, and only the coefficients are refitted during the cross-validation. This will also impact the reported performance, and is one of the topics that will be discussed in the next chapter (Chapter 10).

9.3.3 Challenges and limitations

Most LUR models are developed for single cities and a model built for one location is not necessarily applicable at other locations (Hoek et al., 2008; Knibbs et al., 2014). In the context of LUR studies, *transferability* is used to denote the possibility of transferring a fitted model to other locations and providing reliable results without adjustment of the model. Further, *generalizability* denotes whether general models can be developed for multiple locations at once that are applicable to one location (Jerrett et al., 2005; Hoek et al., 2008; Patton et al., 2015). It is also possible to perform a local recalibration of a transferred model to account for differences in e.g. vehicle fleet composition and background concentrations. Fewer sites are most likely needed in such case (Briggs et al., 2000; Hoek et al., 2008). In general, it is observed that the performance of transferred models is worse than those of site-specific models. In the study of Patton et al. (2015), LUR models were developed for different neighbourhoods of a city. Models directly transferred

from one neighbourhood to another generally performed poorly ($R^2 < 0.17$). Wang et al. (2014) developed a Europe-wide model as part of the ESCAPE study. In some areas of central Europe, the European model performed poorly, but in general the European LUR models were found to have a reasonable power to predict spatial variations in areas not used for model building (a median R^2 of 0.59 for NO_2 and of 0.67 for $PM_{2.5}$ absorbance for the different study areas). According to Hoek et al. (2008), more generalizable regression models could be developed if greater emphasis were placed on developing models with variables applicable to multiple locations than on maximizing the R^2. However, as LUR models are empirical models, in contrast to dispersion models which are derived from physical principles, limitations in transferability may be inevitable to some extent (Hoek et al., 2008).

Another difficulty for transferring LUR models, and for the development of LUR models in general, is the availability and accessibility of geographic datasets and the consistency in the definition and calculation of the geographic variables (Hoek et al., 2008; Li et al., 2012). For example, traffic intensity data are not always readily available for all regions. However, in most LUR models they are very important variables.

Most LUR studies focus on characterizing long-term averaged spatial variation in pollutant concentrations. However, including a temporal component is of interest for studies that need exposure estimates at a finer temporal scale. This will be discussed in more detail in Chapter 11.

9.3.4 LUR models based on mobile monitoring

LUR modelling requires air quality measurements at many locations across the study area to characterize the complex urban environment. However, traditionally, stationary campaigns at a small to moderate number of sites are used. Mobile monitoring is an alternative approach to gather air quality data at a high spatial resolution. In recent years, it is also used as a monitoring approach to gather data to build LUR models. However, we have to distinguish between short-term monitoring and continuous mobile monitoring. In the first case, several studies used a mobile platform to perform short-term measurements at many locations. For example, Larson et al. (2009) aggregated 5 to 13 min of data at a certain location (while still moving but in a cloverleaf pattern). In the study of Ghassoun et al. (2015), they performed measurements using a bicycle and stopped at the predefined sampling spots to conduct a 3 min measurement. Montagne et al. (2015) performed measurements during 30 min per site using an electric vehicle. Deville

Cavellin et al. (2015) collected data at 76 sites, where each site was visited 6 to 12 times for 20 min across three seasons.

In the second case, mobile measurements are performed continuously on-road, without stopping specifically to make a measurement at a specific location. Only few studies use this kind of mobile measurements as a basis for LUR modelling. For example, Hasenfratz et al. (2015) and Mueller et al. (2016) presented a study on the modelling of particle number concentration in Zurich using data from a tram-based mobile sensor network. Hankey and Marshall (2015a) used bicycle-based, mobile measurements to build LUR models, and in studies of Kanaroglou et al. (2013), Patton et al. (2014a) and Weichenthal et al. (2016b) van-based measurements were used. An overview of the studies using continuous mobile measurements to build LUR models is given in Table 9.1.

To build LUR models, mobile monitoring has the advantage of being able to deliver a data set with a much higher density of sampling locations. However, using such data also gives rise to new issues. These include the aggregation of the data and consequences of the shorter time period that measurements are performed at each sampling location. These issues will be explored in practice and discussed in more detail in the next chapter.

Table 9.1: Overview of studies using continuous mobile monitoring as basis for LUR model development.

Reference	Location	Mobile monitoring	Pollutant	Type	Regression
Kanaroglou et al. (2013)	Hamilton, Ontario, Canada	van, equipment mounted on roof	SO_2	spatial	linear
Patton et al. (2014a), Patton et al. (2015)	Somerville, Massachusetts, USA	van	PNC	spatio-temporal	GAM
Hasenfratz et al. (2015), Mueller et al. (2016)	Zurich, Switzerland	tram	PNC	spatial, spatio-temporal	GAM
Hankey and Marshall (2015a)	Minneapolis, Minnesota, USA	bicycle, 3 routes	PNC, BC, $PM_{2.5}$, particle size	spatial	linear
Sabaliauskas et al. (2015)	Toronto, Canada	walking, 10 routes	UFP	spatial	linear
Weichenthal et al. (2016b)	Toronto, Canada	van, equipment mounted on roof, no fixed routes	UFP	spatial	linear
This study	Antwerp, Belgium	bicycle and on foot, no fixed routes	BC	spatial, spatio-temporal	linear, SVR

CHAPTER 10

Development of land use regression models for black carbon based on mobile measurements in the urban environment

10.1 Introduction

Land use regression (LUR) modelling requires air quality measurements at multiple locations across the study area. Typically, stationary monitoring is used at 20 – 100 locations (Hoek et al., 2008). However, Basagaña et al. (2012) argue that LUR models for complex urban settings should be based on a large number of measurement sites (> 80 in their study). Mobile monitoring can provide an alternative way to gather data at a high spatial resolution. Some studies use a mobile platform to perform short-term measurements at many locations more easily (e.g. Larson et al., 2009; Merbitz et al., 2012; Ghassoun et al., 2015; Montagne et al., 2015). Only few studies use mobile measurements as a basis for LUR modelling. For example, Hasenfratz et al. (2015) and Mueller et al. (2016) present a study on the modelling of particle number concentration in Zurich using data from a tram-based mobile sensor network. Hankey and Marshall (2015a) use bicycle-based, mobile measurements to build LUR models, and in studies of Kanaroglou et al. (2013), Patton et al. (2014a) and Weichenthal et al. (2016b) van-based measurements are used. Mobile measurements also have the potential of being collected

in participatory and community-based campaigns. Volunteers can systematically collect targeted data sets, or data are collected opportunistically during (repeated) daily activities or trips, with the potential to provide improved estimates of spatial variability (Snyder et al., 2013, Chapter 7).

In this chapter, we will investigate the development of LUR models based on opportunistic mobile measurements to predict annual average concentrations at a high spatial resolution in the urban environment. This case study is based on measurements gathered by city wardens during their surveillance tasks, which were presented before in Chapter 7. The measurement campaign resulted in a higher spatial density of measurement locations compared to most LUR studies (sampling points at an approximate resolution of 50 m along the roads). Different techniques to build the LUR models, both linear and non-linear, and including different methods to select the relevant predictor variables, will be evaluated. For the evaluation, a custom spatial cross-validation scheme is used, and attention is paid to the implementation of the cross-validation to ensure a proper assessment of the predictive ability of the model.

10.2 Materials and methods

10.2.1 Study location and description

The study site is the city of Antwerp, Belgium, a medium-sized city of 480,000 inhabitants (51°12' N, 4°26' E, 985 inhabitants km^{-2}). The inner city (within the ring road) has an area of approximately 16 km^2. The study area where measurements are gathered comprises a quarter of this region (approximately 3.7 km^2), and is shown in Figure 10.1. This region consists mainly of residential and commercial areas, including main traffic roads and green areas, and a highway (the ring road) is located at the border of the study area. There is no heavy industry located within the study area itself, but the port of Antwerp borders the city at the north. There are no significant differences in elevation throughout the study area.

10.2.2 Mobile air quality monitoring

The opportunistic campaign was carried out with the collaboration of city wardens from July 2012 until June 2013. The Antwerp city wardens are city employees with a surveillance task and are outdoors for a large part of the day carrying out

surveillance tours by bicycle or on foot. These surveillance tours do not follow fixed routes or times. For a more detailed description of the campaign and data quality control and processing, see Section 7.2.1.

10.2.3 Aggregated BC concentrations

As described in Chapter 7, the data at 1 s resolution are aggregated over segments of approximately 50 m resolution along the roads. This results in different passages for each segment, where one passage is a continuous period of time during which measurements are performed in that segment. For each segment, an aggregated concentration level is calculated based on all passages using the trimmed mean and is temporally adjusted to an annual average concentration. These are the data points that will be used as the dependent variable in the LUR models. Because no fixed routes were followed, the number of passages was not the same for all segments. Only those segments with at least 5 passages are used for the model, and this results in 1457 sampling locations.

A few of the segments close to the ring road were removed from the target data set, in particular, the segments located at a bridge over the ring road. These data are not representative for the ring road itself and those high values for the traffic variables were not well represented within the dataset.

10.2.4 GIS data

Data were gathered for four categories of predictor variables: traffic variables (traffic intensity, road length, distance to roads), land use, population density and physical geography (urban morphology). The elevation is not considered as predictor variable. The different data sources are (i) OpenStreetMap (OSM), (ii) Urban Atlas, (iii) Central Reference Address Database (CRAB), (iv) a traffic model, (v) sky view factor data (open data Antwerp) and (vi) data on biking lanes from Province of Antwerp. Based on these data, predictor variables were calculated based on different buffer sizes around the measurement locations or as a point estimate, similarly to previous studies (e.g. Eeftens et al., 2012a; Dons et al., 2013b). An overview of these variables is given in Table 10.1.

OpenStreetMap OpenStreetMap is a collaborative and openly licensed map[1]. In particular, the information on roads (the *highway*-key in the database) is used.

[1] `https://www.openstreetmap.org/about`

This data includes the type of roads (primary, secondary, motorway, etc.) and the distance from the measurement location to the closest road was calculated, as well as the total road length in surrounding buffers.

Urban Atlas The Urban Atlas[2] is providing pan-European comparable land use and land cover data for large urban zones based on satellite images. Land use classes including residential, industrial, port, airport and green urban areas were used.

Address locations The Central Reference Address Database (CRAB) is a freely available[3] database containing street names, house numbers and information about the geographical positioning of addresses for Flanders. The address positions were used to calculate the number of addresses in the surrounding area, which could possibly link the air quality to domestic emissions (e.g. heating).

Traffic data Traffic intensity data was obtained from a traffic model specifically developed for the city of Antwerp (SGS, 2010) and previously applied in a study of Lefebvre et al. (2013). The model provides average daytime traffic intensity for light, medium and heavy traffic, and for all streets in the study area. Different variables were calculated: the total traffic load (sum of light, medium and heavy traffic), the heavy traffic load and the fraction of heavy to total traffic load.

Sky view factor The sky view factor (SVF) is a measure of the total fraction of visible sky from the position of an observer on the ground. This can be used as a street canyon indicator (Eeftens et al., 2013). The SVF can be calculated using 3-dimensional building data or a digital surface model, but for the city of Antwerp the SVF data were made available as open data[4].

Cycling-specific data Thanks to the Province of Antwerp and the Fietsersbond (cyclists association), there is an extensive dataset available describing several aspects of biking lanes, including the distance to the traffic lane. This information was used as an extra variable (distance_to_traffic). Because those data were not available for all streets, they were combined with dist_near (calculated based on the OpenStreetMap data) to fill the missing locations.

[2] http://www.eea.europa.eu/data-and-maps/data/urban-atlas
[3] https://www.agiv.be/international/en/products/crab-en
[4] http://opendata.antwerpen.be/datasets/skyviewfactor-hittekaart

Table 10.1: Overview of the predictor variables calculated from the GIS data.

Name	Description	Source	Buffer radii (m)
res_hd_xx	High-density residential area in a buffer with size XX m (Urban Atlas code 11100 and 11210) [m^2]	Urban Atlas	[100, 300, 500, 1000, 3000, 5000]
res_ld_xx	Low-density residential area in a buffer with size XX m (Urban Atlas code 11120, 11130 and 11240) [m^2]	Urban Atlas	[100, 300, 500, 1000, 3000, 5000]
industry_xx	Industrial or commercial area in a buffer with size XX m (Urban Atlas code 12100) [m^2]	Urban Atlas	[1000, 3000, 5000]
port_xx	Port area in a buffer with size XX m (Urban Atlas code 12300) [m^2]	Urban Atlas	[3000, 5000]
airport_xx	Airport area in a buffer with size XX m (Urban Atlas code 12400) [m^2]	Urban Atlas	[1000, 3000, 5000]
urban_green_xx	Urban green area in a buffer with size XX m (Urban Atlas code 14100) [m^2]	Urban Atlas	[300, 500, 1000, 3000, 5000]
nature_xx	Natural land in a buffer with size XX m (Urban Atlas code 30000 and 40000) [m^2]	Urban Atlas	[300, 500, 1000, 3000, 5000]
address_xx	Number of adresses in a buffer with size XX m	CRAB	[50, 100, 300, 500, 1000, 3000]
trafnear	Traffic intensity on the nearest road [Veh day^{-1}]	Traffic model	-
trafnear_heavy	Heavy traffic intensity on the nearest road [Veh day^{-1}]	Traffic model	-
trafload_xx	Sum of (traffic intensity * road length) in a buffer with size XX m [Veh day^{-1} m]	Traffic model	[50, 100, 300, 500, 1000]
trafload_heavy_xx	Sum of (traffic intensity (heavy traffic) * road length) in a buffer with size XX m [Veh day^{-1} m]	Traffic model	[50, 100, 300, 500, 1000]
trafloadhv_fraction_xx	Fraction of heavy traffic in a buffer with size XX m	Traffic model	[50, 100, 300, 500, 1000]
roadlength_xx	Total road length in a buffer with size XX m [m]	OpenStreetMap	[50, 100, 300, 500, 1000]
roadlength_major_xx	Total major road length in a buffer with size XX m [m]	OpenStreetMap	[50, 100, 300, 500, 1000]
dist_near	Distance to the nearest road [m]	OpenStreetMap	-
dist_near_major	Distance to the nearest major road [m] (OpenStreetMap primary and secondary)	OpenStreetMap	-
dist_highway	Distance to the nearest highway [m]	OpenStreetMap	-
distance_to_traffic	Distance between bike lane and traffic [cm]	Province of Antwerp	-
skyviewfactor	Fraction of visible sky	Open Data Antwerp	-

The relationship between the BC concentrations and a predictor variable, e.g. distance, is often not a linear function. Therefore, some transformations of the variables were added as additional variables (inverse distance and squared inverse distance functions for dist_near, dist_near_major and distance_to_traffic). Further, the following interactions between different predictor variables were added: traffic (trafload_50, trafnear, trafnear_heavy) multiplied with inverse (squared) distance, and sky view factor interaction with traffic intensity. These interactions were added because both the distance and the local geometry (sky view factor) can have an influence on the contribution of traffic to the local concentration levels.

10.2.5 Model building

The objective of LUR modelling is to relate the pollutant concentration with spatial predictor variables in order to be able to estimate the concentration at locations where no measurements are performed. Traditionally, most LUR studies use standard multiple linear regression techniques (Hoek et al., 2008). Some studies use non-linear techniques such as Generalized Additive Models (GAM) (e.g. Hasenfratz et al., 2015; Dekoninck et al., 2015). In this chapter, both multiple linear regression and a non-linear regression technique, support vector regression (SVR) using a radial basis function (RBF) kernel (Smola and Schölkopf, 2004), will be used.

In many papers, the methodology for building LUR models as described in Henderson et al. (2007) and Eeftens et al. (2012a) is used, and which we will also use in this chapter. This is a supervised stepwise forward search of the best subset of predictor variables, based on an optimization of the adjusted R^2 and the significance of the coefficients in the linear regression. The supervised step checks whether the direction of the effect of each predictor variable (i.e. the sign of the coefficient in the linear model) corresponds to the predefined direction expected based on expert knowledge. The predictor variable that gives the highest adjusted R^2 in a univariate regression is used as a starting point. Subsequently, the predictor variable that yields the highest increase in adjusted R^2 is added in a stepwise manner, if the following criteria are met: (i) the increase in adjusted R^2 is greater than 1 %, (ii) all variables have coefficients with a p-value < 0.05 and (iii) the sign of the coefficient agrees with the predefined effect and the sign of the other coefficients in the model does not change.

Next to the aforementioned method (which we will call the 'classic' method), we have chosen to also adopt different approaches of predictor variable selection based on cross-validation. The cross-validation approach is described in more

detail in the next section. In practice, both a forward search limiting the number of variables by requiring an increase in R^2 of 0.01, and a combined forward and backward search of a subset of variables that maximizes the cross-validation R^2 is used. In addition, LASSO, a linear model that estimates sparse coefficients through regularization, is used (Tibshirani, 1996). To summarize, the following model building procedures will be used:

- **No selection**: using all available predictor variables in the model without selecting a subset.
- **Classic** (only for the linear models): using the 'classic' approach of a supervised stepwise forward variable selection.
- **LASSO** (only for the linear models): regularised linear model estimating sparse coefficients (and in this way selecting a subset of predictor variables). This method uses the cross-validation R^2 to optimize the regularisation parameter.
- **Optimal CV**: using a combined forward and backward search for a subset of predictor variables that maximizes the cross-validation R^2.
- **Forward CV**: using a stepwise forward variable selection based on the cross-validation R^2, but with a stopping criterion of an R^2 increase of 0.01.

The SVR models also use cross-validation during model building for all methods to optimize the hyperparameters (the parameters that determine the model structure).

Further, all predictor variables are scaled by subtracting the mean and scaling to unit variance during model building (in case of cross-validation the mean and variance are determined based on the training dataset). For the SVR, a grid search is performed for the optimization of the hyperparameters. All regression analyses are performed using the Python packages scikit-learn (Pedregosa et al., 2011) and Statsmodels (Seabold and Perktold, 2010).

10.2.6 Model evaluation and spatial cross-validation

Performance metrics The LUR models are evaluated using a set of different metrics. First, the R^2 (coefficient of determination) is used. It provides a measure of how well unseen samples are likely to be predicted by the model. If $\hat{y}_i$ is the predicted value and y_i is the corresponding true (measured) value of the i-th

sample (with a total of n samples), then R^2 is defined as:

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\,, \tag{10.1}$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ is the mean true value. In addition to the R^2, also the explained variance (EV) is calculated as:

$$\mathrm{EV} = 1 - \frac{\mathrm{Var}(y - \hat{y})}{\mathrm{Var}(y)}\,, \tag{10.2}$$

where Var denotes the variance. For a linear model with an intercept, the R^2 and explained variance are the same (and equal to the squared Pearson correlation), but once the model is applied on other data (e.g. during cross-validation) or when using SVR, these two metrics are not necessarily the same. There can be a high correlation between the measured and predicted values leading to a good explained variance, but still a poor prediction of the absolute values leading to a low or negative R^2. Further, also the root mean squared error (RMSE) between the model predictions and measurements is calculated.

The air quality measurements at short distances are known to be correlated. The spatial autocorrelation of the BC measurements and of the model residuals is evaluated using an empirical variogram and Moran's I statistic. To calculate the latter, appropriate spatial weights have to be defined. In this study, the inverse squared distance is used for the full matrix of all points.

Spatial cross-validation scheme The metrics above are calculated using cross-validation to provide a more realistic performance evaluation of how well the models would predict the air quality for independent data (e.g. other locations within the same city) compared to the metrics for the model data itself. Given the high spatial autocorrelation in air quality data and given the high spatial density of sampling locations in this study, we have chosen to use a cross-validation scheme based on delineated areas instead of an n-fold cross-validation with random folds. Different zones of $1\,\mathrm{x}\,1\,\mathrm{km}^2$ are constructed within the study area (Figure 10.1). The samples within each zone are used as folds in the n-fold cross-validation ($n = 6$ in this case). At the borders of the zones, there will still be a spatial autocorrelation between samples in the training and test dataset, but regarding the spatial autocorrelation this method is more appropriate than leave-one-out cross-validation (LOOCV) or random n-fold cross-validation. A possible improvement would be not to work with fixed spatial zones, but to use a buffer of certain size around each individual sampling location.

As explained in the previous section, this cross-validation scheme will also be used during the model building (selection of a subset of predictor variables and

Figure 10.1: The different spatial zones for cross-validation. The zones are constructed as 1 x 1 km^2 areas based on the UTM coordinates. Some of the zones with fewer sampling locations are combined into one zone, resulting in six zones as indicated with numbers in the figure.

tuning of hyperparameters). In this scheme, each time the data of one fold (zone) is held out as validation data to estimate the performance of the model that is fitted on the data of all other folds for a given set of predictor variables. In the model building procedure, we can distinguish two steps. Firstly, suitable predictor variables are selected and/or hyperparameters are optimized, which we will jointly refer to as model selection. Secondly, given a certain subset of variables or values for hyperparameters, the model is fitted on the data and the parameters are estimated.

If the purpose is to assess the predictive power of the model on unseen data, it is important that the validation data are not used in the full model building process (both model selection and parameter estimation). Hence, using the same cross-validation (with the same folds) that is used during model selection for performance estimation, can still give an optimistic estimate of the performance. To obtain an unbiased estimate of the predictive performance, a third subset of the data is needed that is not only not used in the parameter estimation, but also not in the model selection (variable selection, tuning of hyperparameters). Therefore, we will report the performance metrics based on two different cross-validation schemes:

- Cross-validation with only refitting the models (**CV1**): in this scheme, the model is only refitted during cross-validation (parameter estimation). The model selection is only done once using all data.
- Cross-validation with rebuilding the models (**CV2**): in this scheme, both model selection and parameter estimation are performed during the cross-validation. This means that 6 different models (in terms of selected variables and values for hyperparameters) are obtained.

In case the model selection itself also uses cross-validation, CV2 implies that a nested cross-validation will be performed. One fold is held out as test dataset, and the model will be developed on the other data using cross-validation with the remaining 5 folds. This will result in 6 different models (each with its own subset of predictor variables or values for hyperparameters) instead of one. The variability between the models will also give an indication of the stability of the model selection. For the classic approach, the variable selection is not based on cross-validation, so for this method CV2 is not a nested cross-validation. When reporting the performance metrics for the cross-validation, the metrics are calculated using the pooled predictions, meaning the combination of the predictions in each fold.

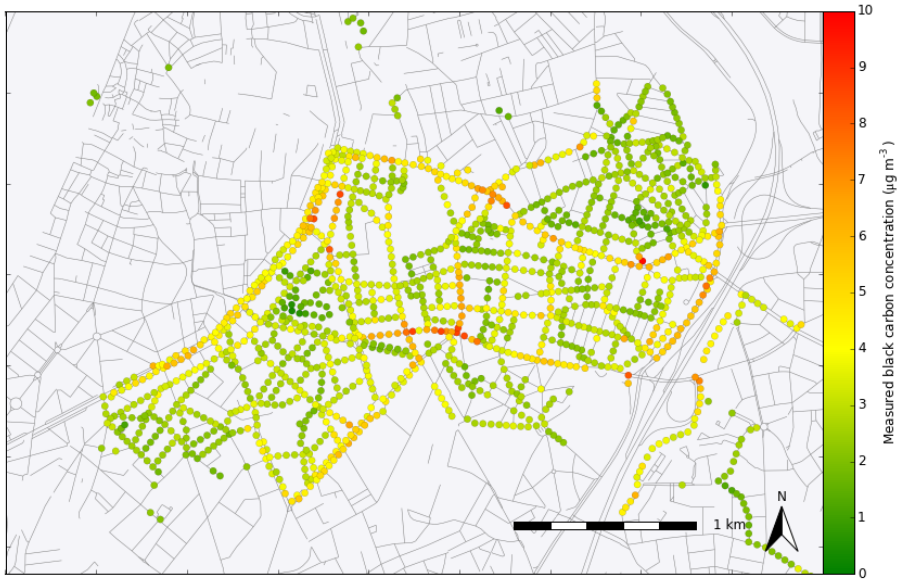


Figure 10.2: Map of the measured BC concentrations.

10.3 Results

10.3.1 Exploration of the target and predictor variables

A map of the measured concentrations is shown in Figure 10.2. The concentrations range from 0.4 μg m^{-3} to 9.8 μg m^{-3} with a mean of 3.3 μg m^{-3} (median of 3.0 μg m^{-3}). The concentrations are spatially correlated. An empirical variogram is constructed to explore how the data are related with distance (Figure 10.3). The largest increase in variance occurs up to 300 m, and after approximately 500 m the BC concentrations are no longer autocorrelated. Moran's I statistic is also calculated and is equal to 0.33 for the measured concentration levels.

Some of the predictor variables were left out (variables related to nature and low density residential areas) because they were not present in the study area itself. Also the largest buffer zones of urban green and high-density residential area (3000 and 5000m) were left out because they had an almost constant value for the study area. The study area only includes urban area, but showed a large variability for many of the other predictor variables.

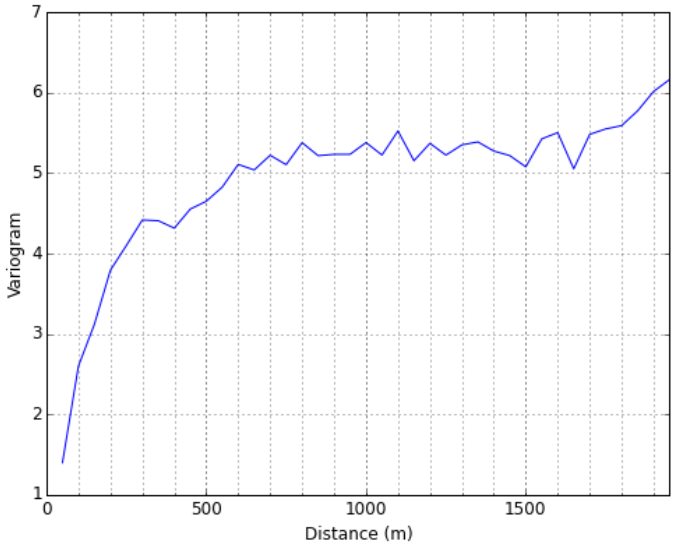


Figure 10.3: Variogram of BC measurements.

10.3.2 Model results

The evaluation results of the different LUR models are summarized in Tables 10.2 and 10.3. For both the linear models and the SVR models, different predictor variable selection methods are used (Table 10.2). Each method is evaluated in different ways (model, CV1 and CV2 performance). First, one model is built based on all data (selection of predictor variables, optimization hyper-parameters), and the model is evaluated using cross-validation (CV1). Second, when using the cross-validation approach where the model is rebuilt during cross-validation, 6 different models are obtained (CV2). For these models, the minimum and maximum of the performance metrics are reported in Table 10.2, as well as the metrics calculated for the pooled predictions of the 6 folds. The selected predictor variables for a selection of the individual models are given in the Supplementary Material. A large variability in selected variables can be noted.

Table 10.2: Overview of the evaluation results of the different LUR models. For the model based on all data, the model R^2 and the cross-validation (CV1) R^2, explained variance (EV) and RMSE are given. Moran's I statistic is calculated for the residuals of the model fitted on all data. For CV2 (cross-validation with full rebuilding), the R^2, EV and RMSE of the different models are given for the pooled predictions and as the minimum and maximum of the individual models as [min, max].

	CV1 (model built using all data)					CV2 (cross-validation with rebuilding)		
	Model R^2	R^2	EV	RMSE	Moran's I	R^2	EV	RMSE
Linear regression								
No selection	0.56	-0.08	-0.07	1.4	0.16	-	-	-
Classic	0.43	0.35	0.35	1.0	0.26	0.24 [-0.65, 0.52]	0.25 [0.28, 0.53]	1.1 [1.0, 1.3]
LASSO	0.35	0.28	0.28	1.1	0.31	0.26 [-0.26, 0.44]	0.26 [0.26, 0.45]	1.1 [1.0, 1.3]
Optimal CV	0.52	0.46	0.46	1.0	0.19	0.22 [-0.70, 0.50]	0.23 [-0.02, 0.52]	1.1 [1.0, 1.3]
Forward CV	0.38	0.34	0.34	1.1	0.29	0.20 [-0.75, 0.36]	0.22 [0.22, 0.41]	1.2 [0.9, 1.3]
SVR								
No selection	0.68	0.3	0.3	1.1	0.13	0.26 [0.06, 0.36]	0.26 [0.16, 0.38]	1.1 [0.9, 1.3]
Optimal CV	0.6	0.46	0.46	1.0	0.16	0.19 [-0.60, 0.40]	0.20 [0.04, 0.43]	1.2 [1.0, 1.3]
Forward CV	0.48	0.4	0.41	1.0	0.23	0.24 [-0.03, 0.37]	0.26 [0.17, 0.41]	1.1 [0.9, 1.3]

Model-building methods

For all methods, the CV2 evaluation results in lower values for the metrics than the evaluation of the models based on all data (model R^2 and CV1 R^2). For the classic linear model, for example, the model R^2 is 0.43 while the CV1 R^2 is 0.35. When rebuilding the model during cross-validation, the score decreases further: a CV2 R^2 of 0.24. A similar trend is found for the other models. Based on the CV2 results, the differences between the different methods are not large. Most of the methods have a similar predictive performance (R^2 ranging between 0.24 and 0.26), with slightly lower values for the forward and optimal CV linear models and the optimal CV SVR model.

For the CV1 R^2, a larger difference between the different models is found. The models that are built by optimizing the R^2 in cross-validation (optimal CV), also score better on the cross-validation (higher CV1 R^2), but the difference between the CV1 and CV2 (nested cross-validation) performance is larger compared to the other models. For example, the linear model based on an optimal subset of variables based on the cross-validation R^2 has a cross-validation CV1 R^2 of 0.46, but only of 0.22 for the nested cross-validation (CV2). This indicates that there is overfitting during the variable selection phase. The cross-validation approach to reduce overfitting during model building is not sufficient, and early stopping to limit the number of variables is still needed. The optimal CV models also have a larger number of variables included: for the linear models, the optimal CV method selects on average 14 variables in CV2 compared to 6 variables for the classic supervised method.

The methods that do not use a custom variable selection method but have regularization built in (the hyperparameters are optimized to prevent overfitting), i.e. LASSO and SVR without manual variable selection, show less difference between the CV1 and CV2 performance metrics. For example, the LASSO model has an R^2 of 0.28 and 0.26 for CV1 and CV2, respectively. For these methods, the cross-validation performed during model building (CV1) gives a good estimate of the predictive performance.

Using the forward CV method for the linear model results in a lower CV2 R^2 compared to the classic method, although the same stopping criterion of a minimum of 1 % increase in R^2 was used for both methods. This indicates that the constraint on a plausible effect of the coefficients of the variables included in the model improves the generalization ability of the model. In the case of SVR, the forward CV selection method gives slightly worse results based on CV2 compared to using all variables (0.24 vs 0.26 for R^2, respectively, and an EV of 0.26 for both). But, the

resulting forward CV model uses only five to six predictor variables in the final models during CV2. Therefore, from a practical point of view (fewer variables are preferred when applying the model), the forward CV method is used for the next paragraph.

Available predictor variables

To investigate the impact of the initial set of available predictor variables used in the model building procedure, both linear and SVR models are built using different initial sets: all predictor variables but without manual transformations and interactions, all predictor variables, all predictor variables without skyviewfactor and distance_to_traffic and all predictor variables without the variables derived from the traffic intensity data (from the traffic model). The following model building procedures are used: the classic approach for the linear models and forward CV for SVR. An overview of the results is given in Table 10.3.

The best results are obtained when using all predictor variables including the transformations and interactions. When not including those, the CV1 R^2 decreases from 0.35 to 0.30 for the linear model. When looking at the selected variables for the best models, in each of the models at least one of trafnear/distance_to_traffic or trafload_50/distance_to_traffic is included (Table 10.S1). Leaving out all variables using skyviewfactor and distance_to_traffic also leads to a lower R^2. The trafnear/distance_to_traffic is replaced by trafnear/dist_near in some cases, but including this variable does not yield the same performance in the cross-validation. Leaving out traffic intensity variables leads to a much lower performance. The traffic variables are mainly replaced by roadlength_major in buffer zones of 50 and 100 m. For the SVR models, similar trends can be noted.

Table 10.3: Overview of the evaluation results of the LUR models starting from different initial subsets of predictor variables. The same metrics are given as in Table 10.2.

	CV1 (model built using all data)					CV2 (cross-validation with rebuilding)		
	Model R^2	R^2	EV	RMSE	Moran's I	R^2	EV	RMSE
Linear regression								
All variables	0.42	0.3	0.31	1.1	0.25	0.19 [-0.78, 0.33]	0.20 [0.20, 0.44]	1.2 [0.9, 1.3]
All variables + interactions	0.43	0.35	0.35	1.0	0.26	0.24 [-0.65, 0.52]	0.25 [0.28, 0.53]	1.1 [1.0, 1.3]
Without skyview/distance_to_traffic	0.42	0.29	0.3	1.1	0.24	0.18 [-0.80, 0.33]	0.19 [0.22, 0.42]	1.2 [0.9, 1.3]
Without traffic	0.31	0.14	0.14	1.2	0.28	0.04 [-1.15, 0.31]	0.05 [-0.08, 0.34]	1.3 [0.9, 1.4]
SVR								
All variables	0.46	0.39	0.39	1.0	0.24	0.20 [-0.53, 0.41]	0.20 [0.07, 0.44]	1.2 [1.0, 1.4]
All variables + interactions	0.48	0.4	0.41	1.0	0.23	0.24 [-0.03, 0.37]	0.26 [0.17, 0.41]	1.1 [0.9, 1.3]
Without skyview/distance_to_traffic	0.56	0.41	0.41	1.0	0.19	0.17 [-0.47, 0.37]	0.18 [-0.03, 0.39]	1.2 [1.0, 1.4]
Without traffic	0.5	0.3	0.31	1.1	0.17	0.05 [-0.48, 0.20]	0.06 [-0.37, 0.21]	1.3 [1.1, 1.4]

Model residuals

The model residuals are analysed by calculating Moran's I value for each of the models (based on the model fitted on all data). For most models there is a small decrease in Moran's I value compared to the value of 0.33 for the measured concentrations. But, the values did not fall to zero, indicating that there still is a considerable spatial autocorrelation in the model residuals. For the classic linear model, Moran's I value decreased to 0.26. The variogram for this model (not shown) shows that the influential distance has become smaller, confirming the decrease in Moran's I. The residuals for this model are also visualised on a map (Figure 10.6).

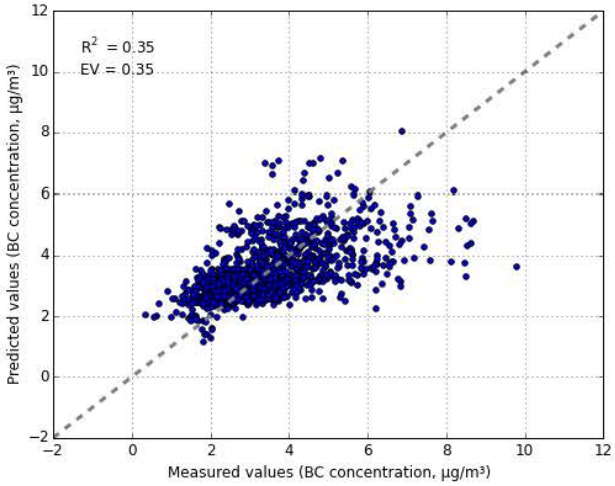
10.4 Discussion

10.4.1 Evaluation of the LUR models

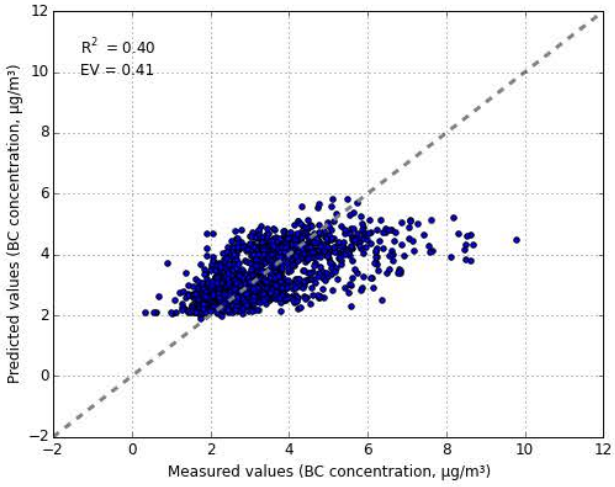
The discussion on the development and evaluation of the LUR models is divided in four paragraphs. Firstly, the different cross-validation approaches to estimate the predictive performance are discussed. Secondly, the different model building techniques are evaluated. Then the actual obtained LUR models and their performance are evaluated and compared to similar studies in the literature, and finally the residuals are discussed.

Different evaluation approaches

The performance of the LUR models was evaluated using a custom spatial 6-fold cross-validation scheme to ensure spatial independence between the training and the test set, given the significant spatial correlation in the BC measurements. Further, the performance was assessed both by cross-validation of the model based on all data (CV1) and by cross-validation with full rebuilding of the model (CV2). Generally, for most models, there is a clear difference between the evaluation approaches resulting in lower R^2 and EV values for CV2 compared to CV1. The CV2 approach is to ensure an unbiased estimate of the predictive performance by excluding the validation data from the model building phase. This indicates that not only the model R^2, but also the cross-validated R^2, does not necessarily reflect the predictive ability of the model. The results stress the importance of a proper evaluation method when assessing the predictive performance of LUR models.



(a)



(b)

Figure 10.4: Scatterplots of the measured and predicted BC concentrations for the optimal (a) linear model (classic) and (b) SVR model (forward CV). The predicted BC concentrations are the result of the model based on all data predicted with cross-validation (CV1).

Figure 10.5: Map of the predicted BC concentrations (at the same locations where measurements have been performed). The predicted values are the result of CV1 cross-validation of the classic linear model based on all data.

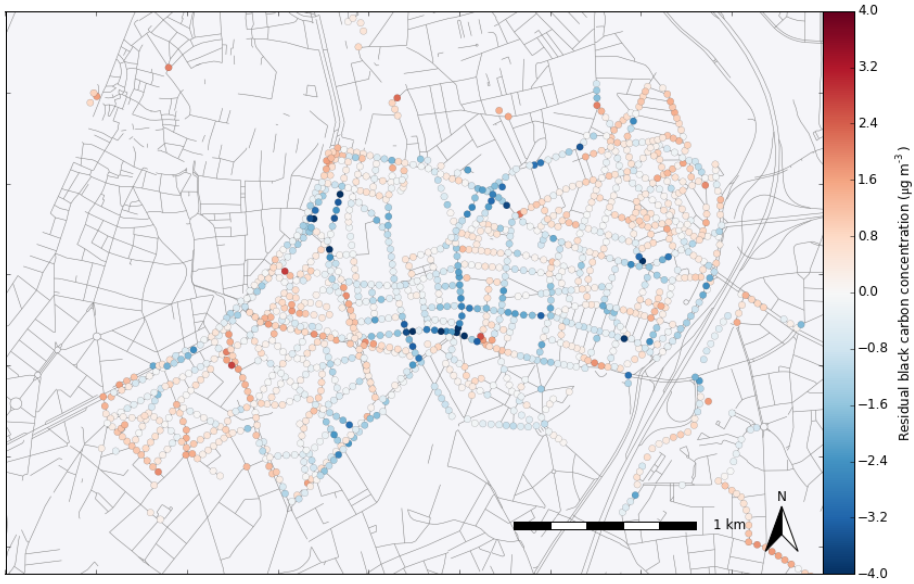


Figure 10.6: Map of the residuals (predicted concentrations minus measured concentrations). Negative values indicate an underestimation of the LUR model. The predicted values are the result of CV1 cross-validation of the classic linear model based on all data.

In literature, LUR models often use leave-one-out cross-validation (LOOCV) to assess the model performance (e.g. Hoek et al., 2008; Eeftens et al., 2012a; Beelen et al., 2013), but it is known that this may overestimate the predictive ability on independent data sets (Wang et al., 2012; Basagaña et al., 2012; Wang et al., 2013). Some other studies used hold-out validation (HV) or random n-fold cross-validation to get a more reliable estimate of the performance. For these approaches, sufficient sampling locations are required to be able to split up the dataset. Using the SVR model without variable selection, the influence of the cross-validation scheme was tested. Using the custom spatial scheme, the R^2 is 0.30 as reported in Table 10.2. When using a 10-fold cross-validation or leave-one-out cross-validation (LOOCV), the R^2 increases to 0.55 and 0.60, respectively.

It is clear that LOOCV-based evaluations are overestimating the predictive ability of the model. However, we have shown that also cross-validation or hold-out approaches can still give an overestimation if the model is not fully rebuilt. Often, the model is not rebuilt but only refitted (parameter estimation), in both LOOCV and HV settings (e.g. Eeftens et al., 2012a; Beelen et al., 2013; Kanaroglou et al., 2013). Our results show the importance of doing a full rebuild (including variable selection) for each fold of the cross-validation to get an estimate of the predictive ability of the model. Because of the high number of sampling locations, we were able to better validate the models (holding out a substantial subset for model evaluation, which is difficult to do with only 20 sampling sites). But at the same time, due to the high spatial density, a more stringent approach (compared to random subsampling) was needed to obtain a spatially independent validation dataset.

Evaluation of the different model building techniques

Different models and model building techniques were tested in this study. In terms of the final predictive ability using cross-validation (CV2), there was not much difference between the different obtained models, although they showed larger differences when looking at the model performance. For the linear models, it is clear that variable selection is important given the negative cross-validation R^2 when using all variables.

Typically, for the variable selection, the supervised stepwise regression is widely used (Hoek et al., 2008; Eeftens et al., 2012a). This approach focuses on selecting models with a limited number of predictor variables that have plausible (predefined) effects. The motivation of this supervised approach is that the model could be applied more easily to other study areas and it also limits the risk of overfitting

(Beelen et al., 2013). In this approach, the variable selection in the model building process is based on all data. Alternatively, the variable selection can also be based on the CV performance, e.g. the cross-validation R^2 instead of the adjusted model R^2, to select those variables that give the best generalization and to minimize the risk of overfitting. Basagaña et al. (2012) compared different techniques (the classic approach with a forward selection based on adjusted R^2, the same algorithm but forward selection based on LOOCV R^2 and the deletion/substitution/addition (DSA, Su et al. 2015a) algorithm that searches through the variable space in order to minimize the squared prediction error during cross-validation). They concluded that the techniques performed similarly in terms of predictive ability on the validation dataset. Our results also show not much difference between the different techniques, and an even worse predictive performance for the models based on a search optimizing the cross-validation R^2 without limiting the number of predictor variables (no constraint on the increase in R^2 for adding a variable). Those models have a better cross-validation R^2 compared to the classic approach, but when applying this technique in the nested cross-validation setting of CV2, the R^2 is lower. This means that, despite the cross-validation during variable selection, there is some overfitting. To have generalizable models, it seems important to limit the number of predictor variables.

The regularized linear model, LASSO, performed slightly better than the classic linear model (CV2 R^2 of 0.26 vs 0.24). The differences are, however, very small and therefore it is not possible to draw clear conclusions. But, LASSO, without supervised variable selection based on predefined effects as in the classic model, did not perform worse. The non-linear SVR model (based on all variables and forward CV) performed similarly when looking at the CV2 R^2 and EV compared to the best linear models. Introducing non-linearities in the model by transforming predictor variables seems to be enough for the linear model to perform as well as the non-linear model. On the other hand, it would maybe not be possible to obtain much higher R^2 values with any technique due to the high variability in the data itself. Weichenthal et al. (2016a) also compared multiple linear regression and a non-linear approach. The non-linear model had a higher model R^2, but when evaluating the model with cross-validation the difference was no longer significant. To conclude, using SVR as a non-linear technique did not yield compelling improvements over linear regression for the data set in this study.

Evaluation of the performance

It is difficult to compare the obtained performance to that of other studies. Firstly, it is not always clear how the model performance is exactly computed. Secondly,

often different methods are used (different metrics or cross-validation approaches) and, as explained above, the choice of the method to evaluate the performance can have a large impact on the obtained values. Further, the number of sampling locations can vary greatly. Based on a resampling exercise, Basagaña et al. (2012) conclude that often a high R^2 is found for LUR models based on a small number of measurement sites, but that this does not necessarily reflect the true predictive ability. For example, in the ESCAPE project, typically high R^2 values are obtained in many of the project's publications (e.g. in Eeftens et al. (2012a), a median model R^2 of 0.89 for PM absorbance is obtained). However, the models and evaluation are based on a low number of sampling locations per study area (20 sites in Eeftens et al. (2012a)). When having a large number of measurement locations, it is more difficult to get predictions with high R^2 values (Basagaña et al., 2012; Hasenfratz et al., 2015). Finally, the lower quality of the data due to the mobile monitoring methodology can also impact the obtained values. Part of the unexplained variance can be attributed to uncertainty in the data itself. We may expect a lower model fit because of an increased variation in concentration (Kanaroglou et al., 2013; Montagne et al., 2015).

In this study, a CV1 R^2 of 0.35 for the classic model is found, although when rebuilding the model during cross-validation (CV2), it decreases to 0.24. For the different models obtained in each iteration of CV2, the validation in one zone yields a negative R^2 (Table 10.S1), but the EV is positive for all models (ranging between 0.28 and 0.53 for the different folds). This means that for this specific zone (zone 1 in Figure 10.1) the absolute values are not predicted well. When comparing to other studies using mobile monitoring, similar performances were found. Hasenfratz et al. (2015) obtained an R^2 of 0.38 for a yearly map with 10-fold cross-validation based on mobile tram measurements. In the study of Kanaroglou et al. (2013), a LUR model for SO2 with a resolution of 50 m was built using mobile van measurements and they obtained an R^2 of 0.30 for a 50 % hold-out cross-validation data set (without rebuilding but with a fixed set of selected variables). Hankey and Marshall (2015a) also used on-road bike measurements, and the LUR model for BC showed an R^2 of 0.20 to 0.35 using a random 1/3 hold-out cross-validation. A higher R^2 of 0.50 based on a validation dataset was found by Weichenthal et al. (2016b) for a UFP model. However, the validation data were scattered throughout the study area (they used those locations with a lower number of data points), and therefore possibly not spatially independent.

The performance during the CV2 cross-validation in this study is lower compared to CV1 and varies considerably between folds: a pooled R^2 of 0.24 is found for the classic linear model, R^2 values for the individual folds of CV2 range between

-0.65 and 0.52. The variability in the selected predictor variables (Tables 10.S1 and 10.S2) and the variation in performance between the different folds during CV2 can be caused by differences in land use between the different zones. This lower performance during CV2 cross-validation indicates that the model still has difficulty to be generalized for the full city and that the performance in predicting the concentrations for locations were no measurements took place (outside of the study area) is limited. In the study of Hankey and Marshall (2015a) mentioned above, they also performed a systematic validation using data from two routes to predict the third route, leading to low R^2 values (0.01 to 0.20 for BC). A possible reason they gave was that the range spanning the predictor variables within each of the routes was not fully balanced. In the study of Patton et al. (2015), measurements were performed in 4 different neighbourhoods. When models built for one of the neighbourhoods were transferred to the other neighbourhoods, the models performed poorly ($R^2 < 0.17$, compared to R^2 of 0.23 to 0.42 for the neighbourhood-specific models). These two studies also have difficulties to generalize the model to other parts of the same city. This evaluation of the transferability of the models is similar to the spatial cross-validation in this study, although the neighbourhoods in Patton et al. (2015) are not contiguous but around 3 to 12 km apart, and the routes in Hankey and Marshall (2015a) also cover a larger area (about 8 x 12 km^2).

Evaluation of the residuals

The model residuals exhibit a considerable spatial autocorrelation. The map of the residuals shows that typically certain streets or part of streets are under- or overestimated in their entirety. This indicates that not yet all explanatory factors are captured in the predictor variables. On the other hand, the spatial autocorrelation in the residuals is also to be expected when using mobile monitoring, because of the short distances between sampling locations. The measured concentrations separated by short distances are likely influenced by similar emission sources and thus any location that is underestimated by the model likely also has neighbours that are underestimated (Hankey and Marshall, 2015a). Further, the model residuals also show a clear relation with the BC concentrations (underestimation of high concentrations and overestimation of low concentrations), and given that the measured BC concentrations themselves are strongly spatially autocorrelated, the residuals will be as well.

Some studies try to remove the spatial autocorrelation in the residuals of the LUR model with an additional model (Li et al., 2012; Kanaroglou et al., 2013). This can be useful for larger areas where regional differences in background concentrations

occur (e.g. the study area in Li et al. (2012) was 160 x 161 km^2). In such case, the non-local variation will be more important and this is not captured by the LUR model. However, this is less relevant when focusing on the smaller scale area of the city centre of Antwerp in this study (no differences in the regional background within the study area).

10.4.2 Selected predictor variables

The traffic intensity variables are the primary selected variables in all models. For example, the SVR model has a CV1 R^2 of 0.29 when only trafload_50 is used as predictor variable, and 0.39 when also including a variable on heavy traffic and the sky view factor. Several of the other variables about address density, industry, airport, port or urban green areas are included in some of the models, but only the traffic variables are included consistently in each model during the CV2 procedure. On the other hand, the models that were forced to not include traffic variables had a low performance. This confirms that traffic is the main factor influencing the local BC concentrations. It also emphasizes the importance of the availability of traffic data when building LUR models. The problem with the traffic intensity variables is that they are not always easily accessible. Hoek et al. (2008) reported that several LUR studies have successfully explored the use of the length of specific road types without traffic intensity data (e.g. Henderson et al., 2007). However, in the present study, the traffic intensity data were essential to obtain models that could predict the BC concentrations.

In the test with different initial sets of predictor variables (Table 10.3), we have also built models for a scenario without the availability of the sky view factor and the cycling-specific bike lane data (distance_to_traffic) because these data are rather specific and not generally available. Given the lower performance of the models, those predictor variables have a clear added value in explaining the variability in BC concentrations in the urban environment. Eeftens et al. (2013) also concluded that street canyon indicators such as the sky view factor could be valuable to consider in air pollution models. The results correspond with the findings in Peters et al. (2014) that traffic intensity, distance to the traffic and street topology are determinant for cyclist exposure.

10.4.3 Mobile monitoring as the basis for LUR models

A limited number of studies have used mobile monitoring to build LUR models, and more research is needed to determine best practices (Hankey and Marshall,

2015a). When using mobile monitoring as the basis for LUR models, an important aspect is how to aggregate the mobile measurements into a suitable form that can be fed into the model. Based on the analysis in Chapter 6, we adopted a 50 m resolution along the roads and the trimmed mean as the aggregation statistic.

Firstly, the data at 1 s resolution were aggregated over a segment of an approximately equal length of 50 m along the road network. In Chapter 6, the effect of different resolutions on the data quality was analysed. They argue that using a higher resolution (e.g. 20 m) results in a higher variability in the data, leading to a need for more repeated measurements, while a lower resolution (e.g. 100 m) brings no significant advantage. Of course, this depends on the number of repeated measurements that are performed, and on the speed of the mobile platform. Hankey and Marshall (2015a) also used this approach, aggregating measurements at equal interval distances along the sampling routes. They tested LUR models based on different spatial resolutions, but found little difference in performance and chose for 100 m. In Patton et al. (2014a), a higher resolution of 20 m was used. Weichenthal et al. (2016b) assigned the collected data to the mid-points of each road segment, which resulted in a mean resolution of 162 m. Another approach instead of aggregating the mobile measurements along the sampling route, is aggregating to a regular grid. For example, Hasenfratz et al. (2015) used a grid of 100 by 100 m and Hasenfratz et al. (2015) a 50 m grid. Such high spatial resolutions are needed to portray the steep concentration gradients found near roads and to minimize exposure misclassification (Batterman et al., 2014).

Secondly, the different passages (aggregated measurement along a 50 m segment) were aggregated to an average concentration at each sampling location using the trimmed mean. This statistic was used to reduce the impact of extreme values on the average concentration of a segment and gives a better estimate of the true mean compared to the arithmetic mean (Chapter 6). Hankey and Marshall (2015a) tested different metrics. They chose the median concentration in their best-case models because this gave a lower error than for the mean concentration. They also showed that lower percentiles yield a better model performance compared to higher percentiles, suggesting that lower percentile (cleaner-air) conditions are more correlated with local land use than higher percentile conditions. The higher percentiles will indeed exhibit a greater variability due to accidental factors. This is partly addressed by using the trimmed mean (or the median in their case). In the following chapter (Chapter 11), we will take a spatio-temporal modelling approach that uses the actual passages and accounts for the temporal variation in the data and thus does not require these decisions.

Mobile monitoring is a suitable method to collect data to build LUR models (Kanaroglou et al., 2013; Patton et al., 2014a; Hankey and Marshall, 2015a). The strength of mobile measurements is the ability to monitor many locations, leading to a high spatial density of sampling locations in the study area. This gives a higher spatial resolution than can be achieved with fixed site monitoring. In this study, 1457 sampling locations were included, and the study of Hankey and Marshall (2015a) used 1101 locations. Basagaña et al. (2012) showed that building LUR models to characterize local air pollution levels in complex urban settings should be based on a larger number of measurement sites than the 20–100 sites that are typically used in LUR studies, and thus the large number of sampling locations obtained by mobile monitoring can yield a more adequate dataset for such LUR models. Mobile monitoring also holds some drawbacks. Due to measuring at many locations for only a short time period, an increased variation in the measured concentrations can be expected. This may result in a lower model fit (Kanaroglou et al., 2013; Montagne et al., 2015). To cope with this, a sufficient number of repeated measurement runs is needed to get a reliable estimate of the average concentration at each location (Chapter 6). Further, mobile measurements will more likely be limited to the day-time hours when people are active and to publicly accessible locations (streets, squares, urban green, etc.).

Dekoninck et al. (2015) present an alternative approach to obtain a city-wide air pollution map using mobile measurements by linking road traffic noise to air quality concentrations. Based on mobile noise measurements, partially combined with BC measurements, a model is built to estimate the yearly average BC concentration from the noise measurements. The model can then be applied at locations were only noise measurements are performed. Mapping noise using mobile measurements would need less repeated measurement to characterize each location compared to mapping BC concentration levels. The noise measurements also add local traffic data including traffic dynamics to the model. Model-based traffic intensity data (as used in this chapter) are more coarse and often do not present the local traffic dynamics at smaller roads very well. But, on the other hand, the advantage is that a LUR model based on those traffic intensity data can also be applied at all locations with similar land use characteristics where no measurements are performed.

10.4.4 Limitations of this study

The applicability of the LUR models obtained in this study is restricted by the characteristics of the input (pollution) data. The measured concentrations are representing street-level (bicycle or pedestrian) daytime exposure values. Further,

the LUR models are applied in a relatively small study area. The question how well the model would perform at a larger scale (e.g. including the peripheries and not only the city centre of Antwerp) remains unanswered. The goal of the present study was to obtain an annual average map and we did not distinguish between different seasons. In some other studies, separate models are built for different seasons, as the influencing factors may vary (e.g. Li et al. (2012) modelled summer and winter seasons separately). But in this case study, we did not have enough data to split the dataset per season.

All conclusions on data quality discussed in Chapter 7 on the dataset of opportunistic measurements used in this study are valid here as well. Although the measurement campaign obtained a high spatial coverage, there is still a rather large uncertainty on the average concentration levels at a spatial resolution of 50 m due to a limited number of measurements for many of the locations and sampling bias. Despite this uncertainty, large spatial patterns within the city are clearly captured with the mobile campaign, which could be modelled with the LUR approach.

10.5 Conclusion

The goal of this chapter was to develop and evaluate LUR models to predict annual average BC concentrations in the urban environment based on opportunistic mobile measurements. We can conclude that mobile monitoring is suited for building LUR models. It can provide the high spatial resolution data needed to characterize the spatial variability in the complex urban environment. It also enables to evaluate the LUR model at a high spatial resolution. The LUR models obtained in this study explain a significant part of the variance in the BC concentrations. The relatively low R^2 values can be partly attributed to the high variability in the data itself due to the mobile monitoring methodology. However, there is still a systematic underestimation of the high concentrations and overestimation of the low concentrations, indicating that not all explanatory factors are already captured in the predictor variables. When using the LUR models to predict the concentrations at locations were no measurements were made, one has to be aware of the potentially low performance in predicting absolute values.

Different model building techniques were tested. LASSO, a regularized linear model, performed slightly better than the classic supervised approach, and the non-linear SVR technique did not show much improvement over a linear model. But, due to the generally low R^2 and the small differences, it is not possible to draw

clear conclusions on which model building technique is preferred. We illustrated the importance of a careful evaluation approach for estimating the predictive performance of the model using an appropriate cross-validation scheme. It is crucial to use independent data (also spatially) for the validation and to not include those test data during variable selection in the model building procedure.

10.A Supplementary data

Table 10.S1: Overview of the selected predictor variables and performance metrics for the classic linear model. Results are shown for the model based on all data and for the different models obtained in each iteration of the spatial cross-validation.

	CV1 (model built using all data)	CV2 (cross-validation with rebuilding) 1	2	3	4	5	6	pooled
Performance								
R^2	0.35	-0.65	0.06	0.3	0.25	0.4	0.52	0.24
EV	0.35	0.28	0.36	0.37	0.28	0.4	0.53	0.25
RMSE	1.05	1.26	1.07	0.95	1.28	1.07	1.03	1.13
Selected features								
address_1000	-	-	-	x	-	-	-	
airport_5000	x	-	x	x	-	x	x	
industry_5000	-	-	-	-	x	-	-	
port_3000	-	-	-	-	-	x	-	
port_5000	-	-	x	-	-	-	x	
res_hd_100	-	x	-	-	-	-	-	
res_hd_1000	x	x	-	-	x	x	x	
trafload_100	x	-	x	x	x	-	x	
trafload_50	x	x	x	x	-	x	x	
trafload_50/distance_to_traffic	-	-	-	-	x	-	-	
trafload_heavy_300	-	x	-	-	-	x	-	
trafnear	-	-	-	-	x	-	-	
trafnear/distance_to_traffic	x	x	x	x	-	x	x	
urban_green_300	x	-	x	-	-	x	x	

Table 10.S2: Overview of the selected features and performance metrics for the forward CV SVR model. Results are shown for the model based on all data and for the different models obtained in each iteration of the spatial cross-validation.

	CV1 (model built using all data)	CV2 (cross-validation with rebuilding) 1	2	3	4	5	6	pooled
Performance								
R^2	0.40	0.19	0.24	-0.03	0.18	0.37	0.24	0.24
EV	0.41	0.26	0.31	0.17	0.23	0.41	0.35	0.26
RMSE	1.00	0.88	0.96	1.15	1.33	1.09	1.29	1.14
Selected features								
address_100	-	-	-	-	-	x	-	
airport_1000	-	-	x	-	-	x	-	
airport_3000	-	-	-	-	-	-	x	
industry_1000	-	-	-	-	x	-	-	
port_3000	-	-	-	x	-	-	-	
roadlength_300	-	-	-	x	-	-	-	
skyviewfactor	x	x	x	-	x	x	-	
trafload_50	x	-	x	x	-	x	-	
trafload_50/skyview	-	x	-	-	x	-	x	
trafload_heavy_300	-	x	-	-	-	-	-	
trafloadhv_fraction_1000	-	-	-	-	-	-	x	
trafloadhv_fraction_300	x	-	x	-	x	x	-	
trafnear	-	-	-	-	x	-	-	
trafnear/dist_near	-	x	-	-	-	-	-	
trafnear/distance_to_traffic	-	-	-	-	-	-	x	
trafnear_heavy/dist_near	x	x	x	-	-	x	-	
trafnear_heavy/distance_to_traffic	-	-	-	x	-	-	x	
urban_green_500	-	-	x	x	-	-	-	

CHAPTER 11

A spatio-temporal land use regression model to assess street-level exposure to black carbon

11.1 Introduction

Land use regression (LUR) modelling is used to assess spatial variation, for example in epidemiological studies to estimate exposure to air pollution (Jerrett et al., 2005; Hoek et al., 2008; Brauer et al., 2008; Beelen et al., 2014). In most cases, the LUR models focus on the spatial variation in annual average concentrations and do not include a temporal dimension. However, in many applications, the temporal variability is an important factor for the exposure. For example, spatio-temporal models can be used in epidemiological studies focusing on an individual-specific period (e.g. pregnancy (Ghosh et al., 2012) or days before death (Maynard et al., 2007)), in short-term exposure studies or to provide personalized exposure information based on GPS tracks.

To incorporate the temporal dimension in LUR models, different approaches are used in literature: temporal adjustment of annual average model output (e.g. Brauer et al., 2008; Wu et al., 2011; de Nazelle et al., 2013), more complex hierarchical (Szpiro et al., 2010; Li et al., 2013) or semiparametric (Gryparis et al., 2014) models, separate models for each typical hour (Dons et al., 2013b) or for each time period (Hasenfratz et al., 2015; Mueller et al., 2016). A temporal LUR

model can also be obtained by including time-dependent data that are related with the temporal variability of the air quality in the model (e.g. Maynard et al., 2007; Patton et al., 2014a; Montagne et al., 2015; Ragettli et al., 2014). Possible time-dependent variables include meteorological variables (wind speed and direction, temperature) and air pollution measurements at fixed monitoring stations.

In this chapter, a spatio-temporal LUR model to predict street-level exposure is developed by incorporating time-dependent variables in addition to the spatial predictor variables. The approach builds further upon the models obtained in the previous chapter, but instead of using the average concentration of all measurements at each location as in the previous chapter, hourly BC concentrations are used as the target to build the LUR model. The model will be validated by predicting the exposure during the trips made by the city wardens and compared with other methods to assess this exposure. The potential to use the opportunistic measurements to develop a dynamic real-time pollution map is discussed.

11.2 Materials and methods

11.2.1 Mobile air quality monitoring

The same dataset of mobile BC measurements as in the previous chapter for the spatial LUR model is used (the opportunistic measurement campaign with the collaboration of city wardens). For a more detailed description of the campaign and data processing, see Section 7.2.1. Three teams of city wardens performed mobile measurements while on surveillance tour, walking or cycling in the city. These surveillance tours do not follow fixed routes or times. The measurements at 1 s resolution are aggregated over segments of approximately 50 m resolution along the roads. This results in several passages, where one passage is a continuous period of time that measurements are performed in a segment. But, instead of aggregating the passages to one average BC concentration for each sampling location (50 m segment) as in the previous chapter, the passages within one hour at the same sampling location are averaged. This results in a dataset where each data point (hourly passage) is characterized by (x, y, t) with t the timestamp with hourly resolution. This resulted in 30,099 hourly data points that are used for building the spatio-temporal LUR model.

For the validation of the model, the data are also aggregated to *trips*. A trip is the path that one team of city wardens followed on their surveillance tour during

one day. For each trip, the average BC concentration is calculated. In total, for the three teams combined, there are 136 trips spread over 110 days.

11.2.2 LUR model and predictor variables

To build the spatio-temporal LUR model, a non-linear regression technique, support vector regression (SVR) using a radial basis function (RBF) kernel (Smola and Schölkopf, 2004), is used. The same spatial predictor variables are used as in the previous chapter (see Table 10.1 for an overview). In addition, time-dependent predictor variables are included in the model. These variables are all obtained from monitoring stations of the Flemish Environmental Agency (VMM), and comprise: wind speed and temperature at one location 5 km north of the study area (station BETM802) and BC concentrations at several stations in the region (stations BETR801, BETR802 (urban traffic), BELAL01 and BETM802 (suburban background), BETN016 and BETN029 (rural background)). The urban traffic stations are located within the study area. All time-dependent variables have an hourly resolution.

All predictor variables are included in the model (the 'SVR no selection' method from the previous chapter). The model itself is evaluated using spatio-temporal cross-validation (see the next section for the different cross-validation schemes). Due to computational constraints, the hyperparameters are optimized once based on all data (using spatial cross-validation), and not for each fold during cross-validation separately. In other words, no nested cross-validation is used. Further, all predictor variables are scaled (by subtracting the mean and scaling to unit variance) during model building. All regression analyses are performed using the Python package scikit-learn (Pedregosa et al., 2011).

11.2.3 Model evaluation: spatio-temporal cross-validation

The following metrics are used to evaluate model performance: coefficient of determination (R^2), explained variance (EV) and root mean squared error (RMSE, in $\mu g\,m^{-3}$) (see Eq. (10.1-10.2)). In the previous chapter, a spatial cross-validation scheme was used to guarantee spatially independent data during model evaluation (given the high spatial autocorrelation in BC concentrations). Here, we also have to deal with the temporal autocorrelation in the BC concentrations. Therefore, the following cross-validation (CV) schemes are distinguished:

- **Spatial cross-validation**: the study area is divided in six zones, and to guarantee spatially independent data, the model is built based on data of

five of the zones and evaluated for the sixth zone (see Figure 10.1 for the zones).

- **Temporal cross-validation**: similarly as for the spatial CV, but now the dataset is divided in time to have temporally independent validation data. In practice, when evaluating the model for the data of day t, the data of day $t-2$ until day $t+2$ are left out to build to model.
- **Spatio-temporal cross-validation**: combination of the spatial and temporal CV. To evaluate the model for the data of one zone of day t, the data of that zone for all days and the data of days $t-2$ to $t+2$ for all zones are left out. This is illustrated in Figure 11.1 (the black and the hatched area is excluded from model building).
- **Minimal cross-validation**: to evaluate the model for the data of one zone of day t, only the data of that spatial zone for days $t-2$ to $t+2$ are excluded from model building, and not the data of all zones during those days and of all days in that zone. In Figure 11.1, the excluded data is represented by the black and doubly hatched area.

The time window of $\pm$ two days is based on a temporal autocorrelation analysis of the measurements at the fixed VMM monitoring station (the mobile measurements itself had no full temporal coverage, and consequently it was not possible to calculate the temporal autocorrelation based on these measurements). The autocorrelation decreases drastically with increasing time lags, and for a time lag of three days the autocorrelation has dropped to the noise level. Therefore, a time window of $\pm$ two days is used to ensure temporally independent data.

11.2.4 Evaluation strategies

Using the cross-validation schemes, the model is evaluated in different ways. In a first step, the model is evaluated at different aggregation levels: the same aggregation level as the one for which the model was built (i.e. hourly passages), a spatial aggregation and a temporal aggregation. For the spatial evaluation, both the predicted and measured passages are aggregated to an annual average map (in the same way as the passages were treated in the previous chapter to obtain the data for the spatial LUR model, i.e. filtered on a minimum of five passages per location and using the trimmed mean). For the temporal evaluation, the predicted and measured passages are now aggregated over all locations for each day.

In a second step, the spatio-temporal LUR model is evaluated on a trip basis. As defined above, a trip is the path that one team of city wardens followed on their

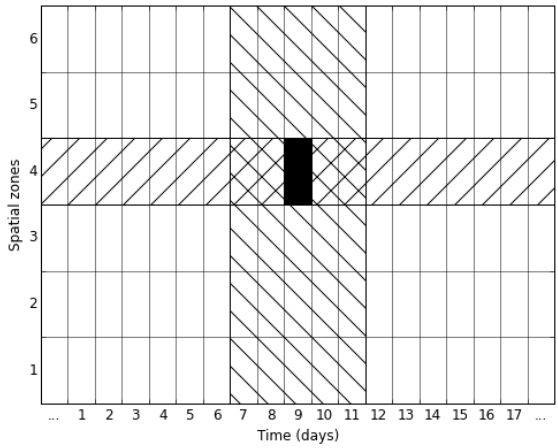


Figure 11.1: Illustration of the spatio-temporal cross-validation scheme: to predict the values of zone 4 and day 9 (black area), all data of the hatched area is excluded from model building. In the minimal cross-validation scheme, only the data of the doubly hatched area is excluded.

surveillance tour during one day. To evaluate the model based on these trips, the hourly BC concentrations for all locations and all time periods are predicted using cross-validation. The predicted average trip exposure is then calculated by combining the predicted concentrations with the GPS track of the trip. Note that the predictions within one trip can come from different folds of the cross-validation and thus from different models.

Next to the spatio-temporal LUR model, the average trip exposure is also estimated with some other methods: (i) based on a fixed monitoring station, (ii) based on a spatial LUR model and (iii) based on a spatial LUR model with temporal adjustment. For the first method, based on a fixed VMM monitoring station, the hourly BC concentration at station BETR801 or BETR802 is used as the city-wide concentration estimate (these stations are located within the study area at a traffic site: BETR802 at the street side and BETR801 30 m away from the street at the same site). For the method based on a spatial LUR, the prediction of the SVR model (no variable selection, spatial cross-validation CV1) obtained in the previous chapter is used to estimate the BC concentration for each location during the trip. Then, a temporal adjustment is used to convert the annual average BC concentration levels of the spatial LUR model to hourly concentration levels.

11.3 Results

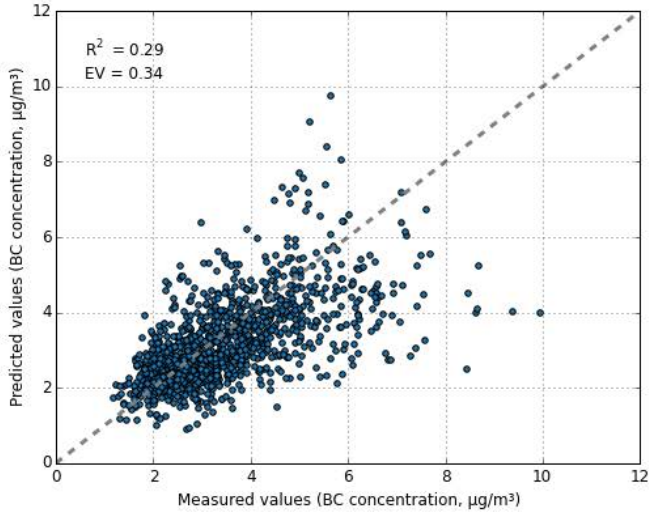
11.3.1 Spatio-temporal LUR model

An hourly spatio-temporal LUR model is built based on hourly passage data. When evaluating the model performance at the same aggregation level as the one for which the model was built, a low performance is obtained: R^2 of 0.15 (EV of 0.18) without cross-validation. When using the spatio-temporal cross-validation scheme, the performance decreases to an R^2 of 0.08 (EV of 0.09). Using the predictions based on the spatio-temporal cross-validation, the model was then evaluated at a spatial and temporal aggregation level. The spatial evaluation yielded an R^2 of 0.29 and EV of 0.34 (Figure 11.2a). The daily averages for the temporal evaluation are shown in Figure 11.2b and this comparison resulted in an R^2 of 0.49 and EV of 0.57. When the BC concentrations at the monitoring stations are left out as predictor variables, the performance of the model is similar based on the spatial evaluation, but the temporal variability is not captured well (R^2 of 0.14, EV of 0.29 for the temporal evaluation).

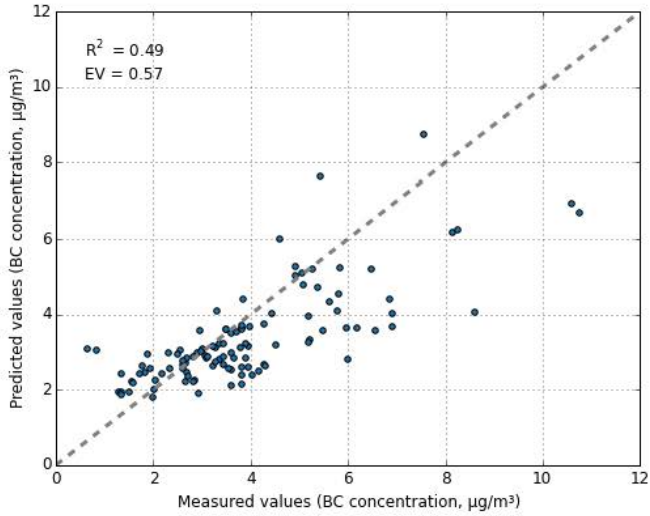
11.3.2 Validation using trips

In a second step, the spatio-temporal LUR model is evaluated on a trip basis. The 136 trips have an average duration of 2.8 h and BC concentration of 3.3 $\mu g\,m^{-3}$, but there is a large variability among the the trips (Figure 11.3). When comparing the predicted and the measured trip exposure, an R^2 of 0.72 is obtained using the minimal spatio-temporal CV scheme (Figure 11.4a). The other cross-validation schemes result in slightly higher or lower R^2 values (Table 11.1).

To benchmark these results, some other methods to estimate the exposure during the trips are implemented for comparison. First, the exposure is estimated based on the measurements from a fixed VMM monitoring station. This resulted in an R^2 of 0.23 (Figure 11.4b) for station BETR801. When using station BETR802 (at the same traffic site but at the street side), the same R^2 but a lower EV (0.34 instead of 0.47) is obtained. Second, a spatial LUR is used. This will only explain the spatial differences in the trip exposure due to the followed route. Using only this spatial LUR, an explained variance of 0.15 is obtained (Figure 11.4c). Third, a temporal adjusted spatial LUR model yielded an R^2 of 0.56 (Figure 11.4d). An overview of the results is given in Table 11.1.

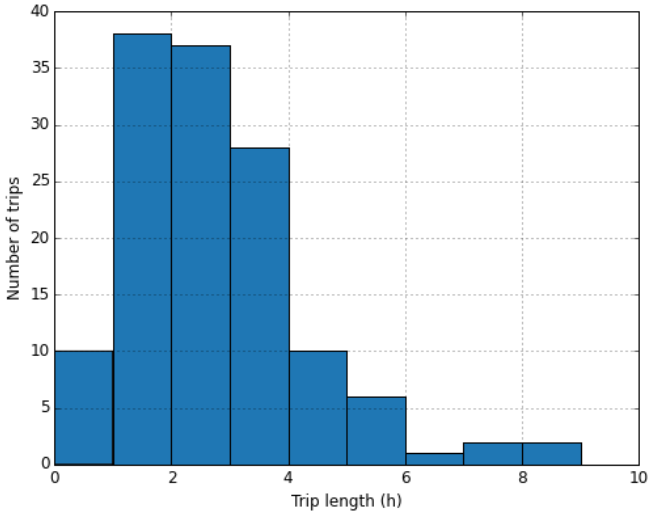


(a)

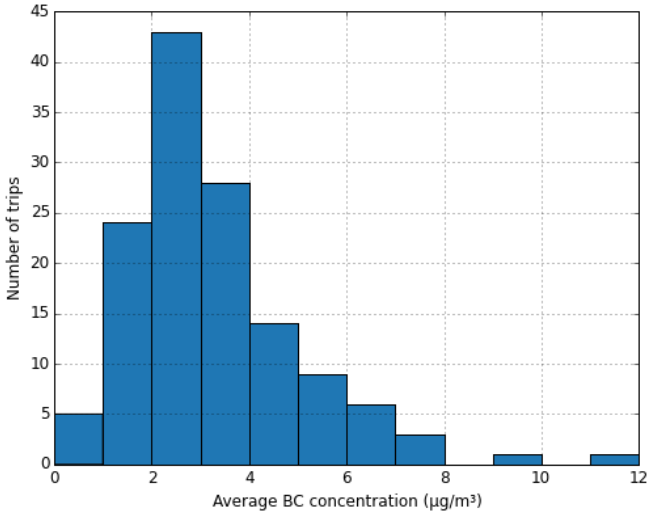


(b)

Figure 11.2: Evaluation of the spatio-temporal LUR model on a (a) spatial level (averages for each sampling location) and (b) temporal level (averages for each day).



(a)



(b)

Figure 11.3: Overview of (a) the length in hours and (b) the average concentration of the trips of the city wardens.

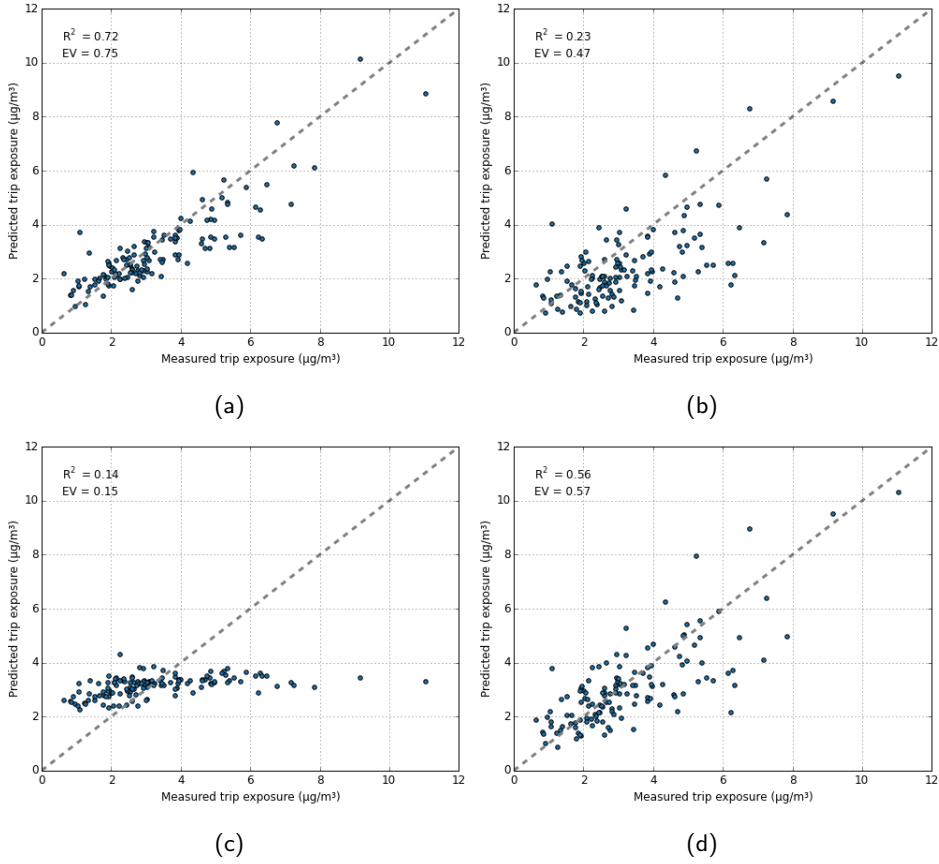


Figure 11.4: Scatterplots of the average concentration during trips based on the measured and predicted BC concentrations for (a) the spatio-temporal LUR model (using the minimal cross-validation), (b) the hourly values at fixed VMM monitoring station, (c) the spatial LUR model and (d) the temporally adjusted spatial LUR model.

Table 11.1: Overview of the evaluation results based on the trips for the different methods and cross-validation schemes.

	R^2	EV	RMSE
Spatio-temporal LUR			
Minimal CV	0.72	0.75	0.9
Spatial CV	0.75	0.75	0.9
Temporal CV	0.58	0.61	1.1
Spatio-temporal CV	0.61	0.61	1.1
VMM hourly values	0.23	0.47	1.5
Spatial LUR model	0.14	0.15	1.6
Spatial LUR model - adjusted	0.56	0.57	1.2

11.4 Discussion

11.4.1 Different approaches for temporal LUR models

Temporal adjustment of annual average concentrations is a frequently used approach to obtain temporally resolved air pollution estimates (e.g. Brauer et al., 2008; Wu et al., 2011; Ghosh et al., 2012; de Nazelle et al., 2013; Dons et al., 2014a). Typically, the temporal variability is introduced by scaling the annual concentrations based on a fixed site monitoring station. When the model is built using temporally integrated data (e.g. passive NO_2 samplers) and thus building a spatio-temporal LUR is not possible, temporal adjustment is a straightforward method to introduce temporal variability. However, this method assumes a stable spatial variability over time. For this reason, Wu et al. (2011) argued that this may not be an improvement or even introduce larger exposure estimation errors than unadjusted (annual average) LUR estimates. Dons et al. (2013b) built separate LUR models for each hour of the day. The hourly LUR models can be considered as annual average models with an hourly temporal resolution, and thus only focus on diurnal variability and are not really spatio-temporal models.

In this study, we included the temporal variability in the LUR modelling approach by using time-dependent variables in combination with the spatial predictor variables. In literature, several studies used the same approach to build spatio-temporal LUR models. Most of these studies conducted repeated short-term measurements at several locations to obtain the data to build the model. For example, Montagne et al. (2015) measured for 30 min at 80 locations in three different seasons. Patton et al. (2014a, 2015), on the other hand, conducted continuous mobile monitoring to build their spatio-temporal LUR model as in this

thesis. Typical variables that are included in these studies are meteorological variables (wind speed and direction, temperature, precipitation, atmospheric stability), time (linear and sinusoidal functions of year, day, and hour), day of the week or indication of weekday/weekend and dynamic traffic intensity (Maynard et al., 2007; Rose et al., 2011; Patton et al., 2014a, 2015). Often, measurements of a fixed site monitoring station are included as well (Rose et al., 2011; Saraswat et al., 2013; Patton et al., 2014a; Ragettli et al., 2014; Montagne et al., 2015). Those measurements can capture the temporal variability in the urban or regional background and include this information in the model. In the study of Patton et al. (2014a), adding central site measurements as a predictor variable instead of meteorological variables did not improve the model. However, they modelled UFP concentrations in close proximity of a highway and wind effects, which were not captured by the central site, were an important factor. Ragettli et al. (2014), on the other hand, found that the concentration at a central station was the most important predictor in their UFP model. Also in the present study, the fixed site BC measurements were essential to model the temporal variability well.

In this study, the measurements of several fixed monitoring stations were included in the model (including both traffic and urban background stations). It should be further investigated what the effect is of including different stations, and how the model would perform if no traffic stations or no stations at all were available in the study area itself. For the spatial variation, the traffic intensity is an important variable, as we have shown in the previous chapter. However, the actual traffic intensity also shows a temporal variability, while the traffic predictor variables used in the models are static (daytime averages). When data from a dynamic traffic model would be available, this could potentially improve the model.

Another approach to build a spatio-temporal LUR model is building separate models for each time period, based only on the measurements performed during that time period. Hasenfratz et al. (2015) and Mueller et al. (2016) built such models based on mobile measurements (using a tram-based sensor network). The problem with this approach, as noted by Hasenfratz et al. (2015), is that it becomes more difficult to obtain accurate models for higher time resolutions, as the number of measurements is reduced. Therefore, they extended their approach by combining the measurements of the modelled time period with measurements from a selection of past time periods based on similarity in meteorological conditions. Mueller et al. (2016) developed models with a high temporal and spatial resolution of 30 min and 10 m by 10 m. They limited predictions in each time period to those locations for which the predictor variables are included in the range of the predictor variables for the locations used to build the model. To increase the area captured by the

model in each time period, the sensor network should be expanded. The approach of instantaneous models can capture the varying spatio-temporal variability better, but it can only be used for periods when real-time measurements are available. In case of the opportunistic campaign in this study, insufficient measurements for each time period were available to use such an approach.

11.4.2 Model performance

The spatio-temporal LUR model shows a low performance when evaluating the model at the same aggregation level as the one for which the model is built, i.e. the hourly passages. However, a good performance should not be expected at this level. The passages do not represent hourly means at the specific location, rather they are a momentary concentration level that can be influenced by many other factors and events (e.g. a passing car) that cannot be included in the predictor variables, and therefore cannot be predicted by the model. To get a more meaningful evaluation, the data have to be aggregated in a certain way to average out part of the inexplicable variance. Therefore, the model is evaluated at a spatial and temporal level. The model is also evaluated based on trips, and this is discussed in the next section.

The spatial evaluation gives comparable results to the spatial LUR model of the previous chapter. Based on the spatio-temporal model, an R^2 of 0.29 and EV of 0.34 is obtained, whereas the spatial LUR model yielded an R^2 and EV of 0.30 (CV1 performance of the SVR model without variable selection, Table 10.2). This confirms that the spatio-temporal model captures the spatial differences in BC concentrations at least as good as the spatial model. The explained variance is even slightly higher. Although the increased complexity of the model (more data points and higher variability in the data points) could lead to a lower R^2, some reasons can explain the comparable performance. First, more data is used to build the spatio-temporal model because no passages are excluded (the criterion to have at least five passages for each sampling location is not imposed here). Second, more predictor variables are included in the model (the time-dependent variables). Part of the variability in the dataset, even after aggregation for the spatial model, is still due to temporal variations in BC concentrations. In the spatio-temporal LUR model, this variation is also modelled (partly). The temporal evaluation shows that the spatio-temporal model captures the larger (daily) temporal variations in the BC concentration levels rather well. Of course, this daily temporal variation is also present in the measured BC concentrations at the monitoring stations which are used as input of the model, so this good performance is expected.

Comparing the model performance obtained in this study to literature is difficult given the specific set-up of the measurement campaign and the evaluation approach. Patton et al. (2014a) also built an hourly LUR model based on mobile measurements (for UFP, using a spatial resolution of 20 m). Validating the model on 1 s resolution, they obtained cross-validation R^2 values of 0.38 to 0.47. However, the reported R^2 values are for the training dataset and not for the validation dataset. It is therefore not possible to compare the R^2 values with the performance obtained in the present study. In the study of Mueller et al. (2016) mentioned above, correlations in the range of 0.67–0.77 are obtained. However, in this evaluation the mobile measurement data were aggregated for 15 min time intervals instead of using the actual data with a resolution of 10 m and 5 s as included in the model. During this time interval, the tram has already travelled a distance of several hundred meters up to a few kilometres. This kind of evaluation is rather comparable to a trip-based evaluation as discussed in the next paragraph.

11.4.3 Predicting street-level exposure during trips

The model is evaluated for its ability to predict the average exposure during trips, making use of the trips of the city wardens. Those trips also form the basis of the data to build the model, but cross-validation is used to ensure the evaluation is performed with independent data. The results show that the spatio-temporal model can predict the trip exposure well (R^2 values between 0.58 and 0.75 depending on the CV scheme). The spatio-temporal LUR model also performs better compared to the other methods. It is clear that the temporal variations in the BC concentration levels are the largest determinant of the trip exposure. The spatial LUR model can only explain 15 % of the variance, while the prediction based on the fixed monitoring station yields an explained variance of 47 % (but gives larger errors leading to a lower R^2 of 0.23). The temporally adjusted spatial LUR model also gives good results, but the spatio-temporal LUR model yields higher R^2 values. The temporal adjustment is constant for the full study area, while the spatio-temporal LUR model can show different temporal profiles at different locations. Dekoninck et al. (2013) used a similar trip-based approach to validate a spatio-temporal model. They predicted cyclist's BC exposure using an instantaneous model based on mobile noise measurements. The average trip exposure was predicted with a correlation of 0.78 (Pearson coefficient). However, this model used simultaneous noise measurements to predict the BC concentration. This has the advantage of including real-time estimates of the traffic intensity. In our model, no measurements are made by the person for whom the trip exposure is predicted and therefore our model is more generally applicable.

Real-time dynamic pollution maps

The spatio-temporal model can also provide a dynamic pollution map. Consider a situation where data are collected continuously in an unstructured way. A model can be constructed that is continuously updated with these new data. The R^2 and EV of the different CV schemes can then be regarded as the predictive ability of the model under different circumstances. The minimal CV scheme gives the predictive ability of such a model at a location where previously measurements were performed, while measurements are performed during the same period at other locations in the city to update the model in real-time. However, as in the case of the city wardens, data collection can be confined to specific areas of the city. For other areas no data will be available at all. The spatial CV scheme then reflects the predictive ability of the model for locations where no measurements were performed at all, but measurements can still be conducted during the same time period at other locations in the city. Data collection can be confined in time as well (e.g. the city wardens did not measure each day). The temporal CV scheme then reflects a situation where no measurements are available for the time period for which the prediction is made, but measurements have been performed during other time periods at the same location. The most stringent scheme, the spatio-temporal CV, relates to a situation where one wants to predict the BC exposure at a location in the city where no measurements were made and at a time when no measurements are performed at other locations. The predictive ability of the model varies for the different CV schemes (R^2 of 0.58 – 0.75, Table 11.1). The highest R^2 values are obtained for the spatial and minimal CV scheme. Thus, adding simultaneous information from similar locations in the other zones increases the predictive ability of the model.

The outlined situations demonstrate the potential of developing a dynamic, real-time pollution map based on measurements that are continuously made throughout the city. By performing continuous but unstructured opportunistic measurements, we are able to dynamically estimate trip exposure with good performance. Potentially, when more devices are deployed and measuring simultaneously or when more measurements are included over time, the model can continue to improve. We have to note that the actual spatio-temporal LUR model in this thesis did not result in a real-time map, as all data of the full measurement campaign were used (not only the past data but also the future data). However, this work demonstrated the potential of the spatio-temporal LUR model as an approach to obtain a real-time map. Further research could explore more intelligent ways to update the model compared to just adding the new measurements to the dataset (e.g. attributing more weight to recent measurements, selective use of historical mea-

surements). It could also be investigated how many measurements are required to obtain a model with good performance. This is, however, beyond the scope of this thesis.

Real-time dynamic pollution maps can provide personalized exposure information. For example, they can be used to estimate exposure of cyclists or pedestrians to traffic-related pollution based on a GPS track, or to find the least polluted route. Some other studies on dynamic pollution maps exist. For example, in the OpenSense project, PNC pollution maps were developed based on a tram-based mobile network (Hasenfratz et al., 2015; Mueller et al., 2016). As explained above, Mueller et al. (2016) modelled PNC at a high spatio-temporal resolution using only the instantaneous measurements (30 min time resolution). Hasenfratz et al. (2015) combined instantaneous measurements with selected historical measurements to improve models with a temporal resolution of 12 h. In the CITI-SENSE project, a data-fusion-based technique is used to combine real-time crowdsourced measurements from a stationary network of low-cost sensors with the output of a dispersion model. This approach allows to produce near-real-time, high-resolution maps of air quality in the urban environment (Schneider et al., 2015). Mobile measurements, however, do not provide a sufficient temporal resolution at one particular location to be suitable for this data fusion approach.

11.5 Conclusion

We demonstrated the use of a spatio-temporal LUR model to estimate the short-term street-level exposure to BC. A spatio-temporal LUR model was built by including meteorological and fixed site BC measurements into the model. The model could capture the spatial variation as well as the spatial LUR model, and the fixed site BC measurements were essential to model the temporal variability. Exposure estimates based on the hourly LUR models were more accurate than based on a fixed site monitoring station and a spatial LUR model, and can be used to estimate exposure of cyclists or pedestrians to traffic related pollution based on a GPS track. We demonstrated the potential of building a real-time dynamic pollution map based on unstructured opportunistic measurements to provide personalized exposure information.

PART IV

EPILOGUE

CHAPTER 12

General conclusions and perspectives

12.1 Urban air quality

Despite the improvements in air pollution levels over the last decades in Europe, air pollution still has a considerable impact on our health and well-being. It leads to many, avoidable premature deaths and an increase in cardiovascular and respiratory diseases. The large majority of the population in Europe is still exposed to PM levels that are considered harmful by the WHO. Air pollution is a factor that affects all people living in a city, independently of their personal life style, and that can be improved with targeted action.

The urban environment shows a high variability in traffic-related air pollutant concentrations. The measurement campaigns in this thesis confirm this, and emphasize the small spatial scale of this variation. In the city of Antwerp, average BC concentrations ranging between 2 and 20 $\mu g\,m^{-3}$ are observed at different locations (at a spatial resolution of 20 m). Large gradients over short distances between different urban micro-environments occur. For example, a 2–3 times increase over a distance of 50 m at the border of an urban green is observed. Also within one street, large differences are measured. The Plantin en Moretuslei has a rather constant traffic intensity throughout the street, but nevertheless shows pollution hotspots that have BC concentration levels 2–3 times higher than the rest of the street.

This high variability indicates the importance of data with a high spatial resolution to get an accurate picture of the pollution levels in a city. Further, it indicates

the importance of local sources and urban planning for the pollution levels of – in this case – BC and UFP. These conclusions can probably be extended to other traffic-related pollutants such as NO_2. At the same time, this large variability also indicates the potential to significantly improve the air quality in cities through local actions.

Pollution maps with a high resolution can be used to visualize and communicate air pollution levels, and to understand the variability and identify possible hotspots. They can be used to estimate the concentration level at the residential address as a proxy for personal exposure in epidemiological research, or be integrated in personal exposure models that include time-activity patterns based on GPS tracks. In this thesis, we have investigated methods to draw pollution maps based on mobile monitoring and on regression models. We have evaluated the use of a spatio-temporal model for outdoor exposure assessment of cyclists or pedestrians.

12.2 Mobile monitoring for air quality mapping

This thesis has evaluated the potential of mobile monitoring to map the urban air quality. We have demonstrated that mobile measurements can provide data with a spatial detail that would not be possible with stationary measurements although maybe not of the same quality at each location. The spatial detail of the observed patterns in the urban air pollution levels clearly shows the added value of mobile monitoring as a complement to stationary monitoring. In this way, hotspots can be identified and the importance of local sources and street geometry can be investigated. The mobile monitoring approach also provides an accessible tool for local measurement campaigns with the participation of volunteers, NGOs or local governments. The airQmap platform and its applications show the potential of this approach. Mobile monitoring will typically not result in a map with full spatial coverage, but it can provide data to build and validate land use regression models or to validate street-level dispersion models.

The spatial resolution of the measurements is increased at the expense of the temporal coverage. This leads to issues on spatial and temporal representativeness that can interfere with the real-life applicability of mobile monitoring. To be representative and useful for personal or community decision making, large quantities of data are required. Mobile measurements have to be repeated regularly and processed carefully. We have developed a set of guidelines on the set-up and data processing of mobile monitoring campaigns, which are summarized in Chapter 8,

and we have showed that by following these guidelines valuable results can be obtained.

Different mobile monitoring approaches can be followed. We have defined a targeted and opportunistic data collection scheme. Targeted, deliberately planned monitoring campaigns can be set up to deliver repeated measurements. This type of monitoring can, for example, be used by authorities or action groups to map a certain area or investigate possible hot-spots. However, this approach is labour intensive compared to stationary measurements. Opportunistic data collection campaigns, on the other hand, take advantage of existing infrastructure or the normal daily routines of commuters or employees. The specific set-up of opportunistic data collection can vary a lot, from unstructured data collection (e.g. city wardens moving throughout the city) to structured data collection following fixed routes. In this thesis we have only explored one specific case. Unstructured data collection leads to some additional challenges and complicates the data interpretation. Due to sampling bias, there is a larger uncertainty on the obtained concentration levels. The targeted campaign (Chapters 5 and 6) resulted in higher quality data, but only a limited number of streets was monitored in this case study. Opportunistic data collection can be a way to gather the large quantities of data that are required for mobile monitoring.

Regardless of the potential, it is important to acknowledge the limitations of mobile monitoring. Even with sound data collection and processing, the lack of temporal coverage is inherent to the mobile monitoring methodology, and will always introduce additional uncertainty. Mobile monitoring will typically result in street-level exposure specific to the used mobile platform. To what extent it can be extrapolated to other transport modes, or to what extent it is representative for residential exposure in a city has, as far as we know, not yet been studied in detail.

Alternatives to mobile monitoring to obtain air quality data at a high spatial resolution exist. A dense network of stationary sensors can also give additional insight. However, the application of such networks is currently still limited by the cost and quality of sensors (see below). When one has access to a limited number of higher quality instruments only, mobile monitoring is a more appropriate approach. Dispersion models are also often used to provide more spatial detail than the fixed monitoring stations. However, models remain an approximation of reality and structural and parametric uncertainties affect the modelled results considerably (Kumar et al., 2011). This is discussed in more detail in the next section.

12.3 High-resolution pollution maps

One of the goals of this thesis was to draw pollution maps with a high spatial resolution. Based on the mobile measurements of the opportunistic campaign, a land use regression (LUR) model was developed. The mobile measurements could capture the small-scale variability, and the question is whether these local patterns can be modelled and thus be predicted at other locations in the city. Based on the exercise in this thesis, we can conclude the model is not yet capable of capturing the local patterns, given the relatively low performance. This can indicate a limitation of the LUR modelling approach to model the small-scale variability within the urban environment. Possibly, this result can also be attributed to the quality of the data. The uncertainty on the data did possibly not allow capturing the more subtle interactions between the different predictor variables. Data from a more intensive mobile monitoring campaign would maybe lead to better results.

Dispersion models are also used to draw pollution maps. However, most dispersion models do currently not predict street-level concentrations, unless they are explicitly coupled with a street canyon module (as is for example done in the study of Lefebvre et al. (2013)). Such models are, however, complex and require detailed input data. This leads to high costs in terms of manpower to deploy the model. To verify whether dispersion models can capture small-scale variability within the city, more validation is needed. Mobile monitoring can deliver the data needed for this. Compared to empirical LUR models, a clear advantage of dispersion models is the ability to perform scenario analyses.

Model-based pollution maps can be used to identify possible hotspots at locations where no measurements were made. However, the spatial LUR model in this thesis can probably not be used for that purpose, as it only captured the broad spatial trends. Traditionally, pollution maps from LUR and dispersion models are also used in epidemiological research to estimate the concentration level at the residential address as a proxy for personal exposure. Whether the street-level LUR models developed based on mobile in-traffic data can be used as outdoor LUR models should be further investigated. LUR models are typically developed based on measurements at the facade, which can be further away from the street. The street-level LUR models also have the potential to be integrated in personal exposure models that include time-activity patterns based on GPS tracks. In such exposure models it could serve as the sub-model for the transport micro-environment. The results of the spatio-temporal LUR model in Chapter 11 showed promising results to model the cyclist exposure based on the GPS track of the trip, and the LUR model performed better compared to an estimate based on the fixed

monitoring station. This technology has the potential to unobtrusively monitor the exposure of large groups of individuals at low cost (de Nazelle et al., 2013). However, the actual usefulness in epidemiological research is still to be investigated (Dons et al., 2014b). Models are inherently limited since peak events will never be modelled. The dominance of the occurrence of peaks in explaining the variation in short-term personal exposure shows the limits of a modelling approach.

This thesis has also paid attention to the validation procedure of LUR models. We illustrated the importance of a careful evaluation approach for estimating the predictive performance of the model using an appropriate cross-validation scheme. It is crucial to use independent data (also spatially) for the validation and to not include those test data during variable selection in the model building procedure. Future LUR studies should more rigorously perform the validation of their models to get a better idea of the predictive ability of LUR models.

12.4 Real-time dynamic pollution maps

In a future scenario, several measurement devices can be constantly measuring throughout the city, carried around by people in their daily routines or attached to moving platforms. These measurements could be used to generate a real-time dynamic pollution map. The spatio-temporal LUR model in this thesis was an exercise in generating such a pollution map based on opportunistic data. The actual model did not result in a real-time map, as all data of the full measurement campaign were used, but it demonstrated the potential of the spatio-temporal LUR model as a first step towards a real-time map. Based on the opportunistic campaign, a good performance in estimating trip exposure was obtained.

Such real-time pollution maps can be used for providing personalized information about air quality to citizens. A potential application is to estimate exposure of cyclists or pedestrians to traffic-related pollution based on a GPS track, or to find the least polluted route. Many challenges still have to be tackled. Further research could explore more intelligent ways to update the model by combining real-time measurements with historic measurements in the model. Another topic is how to deal with data of different sources (mobile and stationary, structured and unstructured, different sensors, model output), typically with different characteristics and of different quality, and to combine them into one pollution map.

The methods to draw pollution maps presented in this thesis can be implemented in different ways. A limited number of measuring devices can already be used in targeted or opportunistic mobile monitoring campaigns, possibly with the partic-

ipation of volunteers, city services or other organizations. This approach provides possibilities for interested parties to implement mobile monitoring campaigns at the present time. The availability of low-cost sensors would enable a larger-scale and possibly participatory deployment of sensors.

12.5 Low-cost sensing devices

In this thesis, the micro-aethalometer has been used as the instrument for conducting mobile measurements. This is an easy-to-use, relatively reliable instrument. Further improvements are desirable, including a better stability regarding temperature and humidity changes, increased battery lifetime, higher flow rate for a lower uncertainty, and methods to correct for filter loading. But, the main hurdle for a more widespread use is the cost of the instrument, which makes it unsuitable for large-scale opportunistic or participatory campaigns.

The choice to focus on the micro-aethalometer in this thesis was made during this PhD research. At the start of the research, four years ago, it was the idea to both use a relatively high-quality but more expensive device, such as the micro-aethalometer, and low-cost gas sensors. More specifically, as part of the EveryAware project, a sensor box was built based on commercially available low-cost gas sensors at the time. However, during the first use cases, it was observed that the performance was not good enough to conduct high-quality mobile measurements. Therefore, we chose to focus on the use of the micro-aethalometer in this thesis. At this point in time, the question can be raised how things have changed. Despite the progress in sensors and sensing devices, we would probably still make the same choice given the context of the project. Nonetheless, sensors have improved a lot over the last years, and a similar exercise as in the EveryAware project, i.e. combining a sensor array and field calibration, would probably already give more promising results.

The use of low-cost sensors would enable larger-scale measurement campaigns. It would make it possible to set up uncoordinated, opportunistic data collection that can deliver the high amounts of data needed for obtaining representative maps with high spatial coverage or for building real-time dynamic pollution maps. However, the current generation of commercially available low-cost sensors is not yet readily usable for ambient air quality monitoring. Some encouraging examples show the potential, but knowledge of the sensing principles and electronics design, and experience in complex data processing and field calibration is still required to yield valuable results. This poses a high barrier to use such sensors. Recently,

some commercial, integrated devices that include strategies to increase performance became available. But, these devices are more expensive, and no rigorous validation is publicly available. However, the technology is advancing rapidly and the availability of usable sensors for ambient air quality monitoring will possibly be a matter of several years (EPA, 2013; Theunis et al., 2016a).

Using low-cost sensors for mobile monitoring introduces additional challenges. More research is needed in this field to investigate the validity of field calibration in a changing environment and the performance at a high temporal resolution. Up to now, we are not aware of a validation of this use case, as most comparisons with references devices are stationary. Given the above, low-cost sensors probably have more potential for stationary applications in the short-term. Nevertheless, also in a dense network of static sensors, the field calibration is a challenging task.

12.6 Citizen science and participatory monitoring

There is an increasing awareness and public interest around air quality issues and a healthy environment in general, and this also results in an increasing active participation. The multiple citizen science projects that have been conducted or started over the last years illustrate this trend. However, the emergence of citizen science projects is hindered by the available instruments. The will of communities to measure is ahead of the availability of readily usable low-cost sensors and the full potential interest in citizen science projects cannot yet be fulfilled. For example, the Air Quality Egg and AirCasting projects have both already built an extensive toolkit including sensing devices, data communication and visualisation platforms, smartphone applications, etc., while the actual sensors they use deliver data of poor quality. For this reason, Ringland plans to use passive samplers in the CurieuzeNeuzen project. This is a long-established, often-used and low-cost method to measure NO_2, but with the disadvantage that it does not give data in real time and cannot capture temporal variations (or only at a very low resolution if the tubes are replaces every few weeks).

Many citizen science projects focus on a stationary network of sensors due to the limitations of the sensors available. Higher quality devices such as the microaethalometer are not suitable for large-scale projects or bottom-up citizen science projects due to the cost of the device. But, such devices are already useful for participatory projects in close collaboration with scientific institutions as the portable monitor in personal and mobile monitoring studies using a limited number of in-

struments. Current examples of this approach are the opportunistic campaign with the participation of the city wardens, personal monitoring studies, and applications using the airQmap platform involving volunteers. In the future, low-cost sensors will increase the opportunities to conduct mobile measurements in larger-scale participatory campaigns and citizen science projects and to collaboratively map the urban air quality.

Bibliography

Al-Dabbous, A. N. and Kumar, P. (2014). The influence of roadside vegetation barriers on airborne nanoparticles and pedestrians exposure under varying wind conditions. *Atmospheric Environment*, 90:113–124.

Aleixandre, M. and Gerboles, M. (2012). Review of small commercial sensors for indicative monitoring of ambient gas. *Chemical Engineering Transactions*, 30:169–174.

Aoki, P. M., Honicky, R. J., Mainwaring, A., Myers, C., Paulos, E., Subramanian, S., and Woodruff, A. (2009). A vehicle for research : using street sweepers to explore the landscape of environmental community action. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2009)*, pages 375–384.

Apte, J. S., Kirchstetter, T. W., Reich, A. H., Deshpande, S. J., Kaushik, G., Chel, A., Marshall, J. D., and Nazaroff, W. W. (2011). Concentrations of fine, ultrafine, and black carbon particles in auto-rickshaws in New Delhi, India. *Atmospheric Environment*, 45(26):4470–4480.

Apte, J. S., Marshall, J. D., Cohen, A. J., and Brauer, M. (2015). Addressing global mortality from ambient PM2.5. *Environmental Science & Technology*, 49(13):8057–66.

Argyropoulos, G., Samara, C., Voutsa, D., Kouras, A., Manoli, E., Voliotis, A., Tsakis, A., Chasapidis, L., Konstandopoulos, A., and Eleftheriadis, K. (2016). Concentration levels and source apportionment of ultrafine particles in road microenvironments. *Atmospheric Environment*.

Baldwin, N., Gilani, O., Raja, S., Batterman, S., Ganguly, R., Hopke, P., Berrocal, V., Robins, T., and Hoogterp, S. (2015). Factors affecting pollutant concentrations in the near-road environment. *Atmospheric Environment*, 115:223–235.

Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J., and Künzli, N. (2012). Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmospheric Environment*, 54:634–642.

Batterman, S., Chambliss, S., and Isakov, V. (2014). Spatial resolution requirements for traffic-related air pollutant exposure evaluations. *Atmospheric Environment*, 94:518–528.

Beckerman, B. S., Jerrett, M., Martin, R. V., van Donkelaar, A., Ross, Z., and Burnett, R. T. (2013). Application of the Deletion/Substitution/Addition algorithm to selecting Land Use Regression models for interpolating air pollution measurements in California. *Atmospheric Environment*, 77:172–177.

Beckx, C., Panis, L. I., Arentze, T., Janssens, D., Torfs, R., Broekx, S., and Wets, G. (2009). A dynamic activity-based population modelling approach to evaluate exposure to air pollution: Methods and application to a Dutch urban area. *Environmental Impact Assessment Review*, 29(3):179 – 185.

Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., Briggs, D. J., and de Hoogh, K. (2009). Mapping of background air pollution at a fine spatial scale across the European Union. *Science of the Total Environment*, 407(6):1852–67.

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.-Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., Birk, M., Cyrys, J., von Klot, S., Nádor, G., Varró, M. J., Ddel, A., Gražulevičien, R., Mölter, A., Lindley, S., Madsen, C., Cesaroni, G., Ranzi, A., Badaloni, C., Hoffmann, B., Nonnemacher, M., Krämer, U., Kuhlbusch, T., Cirach, M., de Nazelle, A., Nieuwenhuijsen, M., Bellander, T., Korek, M., Olsson, D., Strömgren, M., Dons, E., Jerrett, M., Fischer, P., Wang, M., Brunekreef, B., and de Hoogh, K. (2013). Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe the ESCAPE project. *Atmospheric Environment*, 72:10–23.

Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z. J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W. W., Katsouyanni, K., Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A. S., Lanki, T., Turunen, A., Oftedal, B., Nystad, W., Nafstad, P., De Faire, U., Pedersen, N. L., Östenson, C.-G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G. G., Eriksen, K. T., Overvad, K., Ellermann, T., Eeftens, M., Peeters, P. H., Meliefste, K., Wang, M., Bueno-de Mesquita, B., Sugiri, D., Krämer, U., Heinrich, J., De Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-Hensch, N., Künzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M.-Y. Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Katsoulis, M., Trichopoulou, A., Brunekreef, B., and Hoek, G. (2014). Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *The Lancet*, 383(9919):785–795.

Beelen, R., Voogt, M., Duyzer, J., Zandveld, P., and Hoek, G. (2010). Comparison of the performances of land use regression modelling and dispersion modelling in estimating small-scale variations in long-term air pollution concentrations in a Dutch urban area. *Atmospheric Environment*, 44(36):4614–4621.

Bekö, G., Kjeldsen, B. U., Olsen, Y., Schipperijn, J., Wierzbicka, A., Karottki, D. G., Toftum, J., Loft, S., and Clausen, G. (2015). Contribution of various microenvironments to the daily personal exposure to ultrafine particles: Personal monitoring coupled with GPS tracking. *Atmospheric Environment*, 110:122–129.

Berghmans, P., Bleux, N., Int Panis, L., Mishra, V. K., Torfs, R., and Van Poppel, M. (2009). Exposure assessment of a cyclist to PM10 and ultrafine particles. *Science of the Total Environment*, 407(4):1286–98.

Bigazzi, A. Y. and Figliozzi, M. A. (2014). Review of urban bicyclists' intake and uptake of traffic-related air pollution. *Transport Reviews*, 34(2):221–245.

Bond, T. C., Anderson, T. L., and Campbell, D. (1999). Calibration and intercomparison of filter-based measurements of visible light absorption by aerosols. *Aerosol Science and Technology*, 30(6):582–600.

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P. K., Sarofim, M. C., Schultz, M. G., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S. K., Hopke, P. K., Jacobson, M. Z., Kaiser, J. W., Klimont, Z., Lohmann, U., Schwarz, J. P., Shindell, D., Storelvmo, T., Warren, S. G., and Zender, C. S. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552.

Brantley, H. L., Hagler, G. S. W., J. Deshmukh, P., and Baldauf, R. W. (2014a). Field assessment of the effects of roadside vegetation on near-road black carbon and particulate matter. *Science of the Total Environment*, 468:120–129.

Brantley, H. L., Hagler, G. S. W., Kimbrough, E. S., Williams, R. W., Mukerjee, S., and Neas, L. M. (2014b). Mobile air monitoring data-processing strategies and effects on spatial air pollution trends. *Atmospheric Measurement Techniques*, 7(7):2169–2183.

Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R. V., Dentener, F., van Dingenen, R., Estep, K., Amini, H., Apte, J. S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P. K., Knibbs, L. D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J. L. T., Shaddick, G., Anderson, H. R., Vos, T., Forouzanfar, M. H., Burnett, R. T., and Cohen, A. (2016). Ambient air pollution exposure estimation for the global burden of disease 2013. *Environmental Science & Technology*, 50(1):79–88.

Brauer, M., Lencar, C., Tamburic, L., Koehoorn, M., Demers, P., and Karr, C. (2008). A cohort study of traffic-related air pollution impacts on birth outcomes. *Environmental Health Perspectives*, 116(5):680–6.

Briggs, D. (2005). The role of GIS: coping with space (and time) in air pollution exposure assessment. *Journal of Toxicology and Environmental Health. Part A: Current Issues*, 68(13-14):1243–61.

Briggs, D. J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., Van Reeuwijk, H., Smallbone, K., and Van Der Veen, A. (1997). Mapping urban air pollution using GIS: a regression-based approach. *International Journal of Geographical Information Science*, 11(7):699–718.

Briggs, D. J., de Hoogh, C., Gulliver, J., Wills, J., Elliott, P., Kingham, S., and Smallbone, K. (2000). A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of the Total Environment*, 253(1-3):151–67.

Broich, A. V., Gerharz, L. E., and Klemm, O. (2012). Personal monitoring of exposure to particulate matter with a high temporal resolution. *Environmental Science and Pollution Research*, 19(7):2959–2972.

Buonanno, G., Stabile, L., Morawska, L., and Russi, A. (2013). Children exposure assessment to ultrafine particles and black carbon: The role of transport and cooking activities. *Atmospheric Environment*, 79:53–58.

Burke, J., Estrin, D., Hansen, M., Ramanathan, N., Reddy, S., and Srivastava, M. B. (2006). Participatory sensing. In *In: Proceedings of the Workshop on World-Sensor-Web (WSW'06): Mobile Device Centric Sensor Networks and Applications*, pages 117–134.

Cai, J., Yan, B., Ross, J., Zhang, D., Kinney, P. L., Perzanowski, M. S., Jung, K. H., Miller, R., and Chillrud, S. N. (2014). Validation of microAeth® as a black carbon monitor for fixed-site measurement and optimization for personal exposure characterization. *Aerosol and Air Quality Research*, 14(1):1–9.

Campbell, A. T., Eisenman, S. B., Lane, N. D., Miluzzo, E., Peterson, R. A., Lu, H., Zheng, X., Musolesi, M., Fodor, K., and Ahn, G.-S. (2008). The rise of people-centric sensing. *IEEE Internet Computing*, 12(4):12–21.

Can, A., Dekoninck, L., and Botteldooren, D. (2014). Measurement network for urban noise assessment: Comparison of mobile measurements and spatial interpolation approaches. *Applied Acoustics*, 83:32–39.

Cao, J.-J., Zhu, C.-S., Chow, J. C., Watson, J. G., Han, Y.-M., Wang, G.-h., Shen, Z.-x., and An, Z.-S. (2009). Black carbon relationships with emissions and meteorology in Xi'an, China. *Atmospheric Research*, 94(2):194–202.

Castell, N., Kobernus, M., Liu, H.-Y., Schneider, P., Lahoz, W., Berre, A. J., and Noll, J. (2014). Mobile technologies and services for environmental monitoring: The Citi-Sense-MOB approach. *Urban Climate*.

Charron, A. and Harrison, R. M. (2003). Primary particle formation from vehicle emissions during exhaust dilution in the roadside atmosphere. *Atmospheric Environment*, 37(29):4109–4119.

Cheng, Y.-H. and Lin, M.-H. (2013). Real-time performance of the microAeth® AE51 and the effects of aerosol loading on its measurement results at a traffic site. *Aerosol and Air Quality Research*, 13:1853–1863.

Choi, W., He, M., Barbesant, V., Kozawa, K. H., Mara, S., Winer, A. M., and Paulson, S. E. (2012). Prevalence of wide area impacts downwind of freeways under pre-sunrise stable atmospheric conditions. *Atmospheric Environment*, 62:318–327.

Collaud Coen, M., Weingartner, E., Apituley, A., Ceburnis, D., Fierz-Schmidhauser, R., Flentje, H., Henzing, J. S., Jennings, S. G., Moerman, M., Petzold, A., Schmid, O., and Baltensperger, U. (2010). Minimizing light absorption measurement artifacts of the Aethalometer: evaluation of five correction algorithms. *Atmospheric Measurement Techniques*, 3(2):457–474.

Conrad, C. C. and Hilchey, K. G. (2011). A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental monitoring and assessment*, 176(1-4):273–91.

Crouse, D. L., Peters, P. A., van Donkelaar, A., Goldberg, M. S., Villeneuve, P. J., Brion, O., Khan, S., Atari, D. O., Jerrett, M., Pope, C. A., Brauer, M., Brook, J. R., Martin, R. V., Stieb, D., and Burnett, R. T. (2012). Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: A canadian national-level cohort study. *Environmental Health Perspectives*, 120(5):708–714.

Cyrys, J., Eeftens, M., Heinrich, J., Ampe, C., Armengaud, A., Beelen, R., Bellander, T., Beregszaszi, T., Birk, M., Cesaroni, G., Cirach, M., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Ddel, A., Dimakopoulou, K., Eriksen, K., Galassi, C., Gražulevičien, R., Grivas, G., Gruzieva, O., Gustafsson, A. H., Hoffmann, B., Iakovides, M., Ineichen, A., Krämer, U., Lanki, T., Lozano, P., Madsen, C., Meliefste, K., Modig, L., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Nonnemacher, M., Oldenwening, M., Peters, A., Pontet, S., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Sugiri, D., Stephanou, E. G., Taimisto, P., Tsai, M.-Y., Vaskövi, É., Villani, S., Wang, M., Brunekreef, B., and Hoek, G. (2012). Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. *Atmospheric Environment*, 62:374–390.

Cyrys, J., Hochadel, M., Gehring, U., Hoek, G., Diegmann, V., Brunekreef, B., and Heinrich, J. (2005). GIS-based estimation of exposure to particulate matter and NO2 in an

urban area: stochastic versus dispersion modeling. *Environmental health perspectives*, 113(8):987–92.

de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M. J., Badaloni, C., Beelen, R., Bolignano, A., Cesaroni, G., Pradas, M. C., Cyrys, J., Douros, J., Eeftens, M., Forastiere, F., Forsberg, B., Fuks, K., Gehring, U., Gryparis, A., Gulliver, J., Hansell, A. L., Hoffmann, B., Johansson, C., Jonkers, S., Kangas, L., Katsouyanni, K., Künzli, N., Lanki, T., Memmesheimer, M., Moussiopoulos, N., Modig, L., Pershagen, G., Probst-Hensch, N., Schindler, C., Schikowski, T., Sugiri, D., Teixidó, O., Tsai, M.-Y., Yli-Tuomi, T., Brunekreef, B., Hoek, G., and Bellander, T. (2014). Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environment International*, 73:382–92.

de Nazelle, A., Fruin, S. A., Westerdahl, D., Martinez, D., Ripoll, A., Kubesch, N., and Nieuwenhuijsen, M. (2012). A travel mode comparison of commuters' exposures to air pollutants in Barcelona. *Atmospheric Environment*, 59:151–159.

de Nazelle, A., Seto, E., Donaire-Gonzalez, D., Mendez, M., Matamala, J., Nieuwenhuijsen, M. J., and Jerrett, M. (2013). Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environmental Pollution*, 176:92–99.

De Vito, S., Massera, E., Piga, M., Martinotto, L., Francia, G. D., and Di Francia, G. (2008). On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario. *Sensors and Actuators B: Chemical*, 129(2):750–757.

De Vito, S., Piga, M., Martinotto, L., and Di Francia, G. (2009). CO, NO2 and NOx urban pollution monitoring with on-field calibrated electronic nose by automatic bayesian regularization. *Sensors and Actuators B: Chemical*, 143(1):182–191.

Dekoninck, L. (2015). *Spatiotemporal modelling of personal exposure to traffic related particulate matter using noise as a proxy*. PhD thesis, Ghent University.

Dekoninck, L., Botteldooren, D., and Int Panis, L. (2013). An instantaneous spatiotemporal model to predict a bicyclist's Black Carbon exposure based on mobile noise measurements. *Atmospheric Environment*, 79:623–631.

Dekoninck, L., Botteldooren, D., and Int Panis, L. (2015). Using city-wide mobile noise assessments to estimate bicycle trip annual exposure to Black Carbon. *Environment International*, 83:192–201.

Delgado-Saborit, J. M. (2012). Use of real-time sensors to characterise human exposures to combustion related pollutants. *Journal of Environmental Monitoring*, 14(7):1824–37.

Deville Cavellin, L., Weichenthal, S., Tack, R., Ragettli, M. S., Smargiassi, A., and Hatzopoulou, M. (2015). Investigating the use of portable air pollution sensors to

capture the spatial variability of traffic related air pollution. *Environmental Science & Technology*, 50(1):313–320.

D'Hondt, E., Stevens, M., and Jacobs, A. (2013). Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring. *Pervasive and Mobile Computing*, 9(5):681–694.

Dons, E. (2013). *Air pollution exposure assessment through personal monitoring and activity-based modeling*. PhD thesis, Hasselt University.

Dons, E., Int Panis, L., Van Poppel, M., Theunis, J., and Wets, G. (2012). Personal exposure to Black Carbon in transport microenvironments. *Atmospheric Environment*, 55:392–398.

Dons, E., Int Panis, L., Van Poppel, M., Theunis, J., Willems, H., Torfs, R., and Wets, G. (2011). Impact of time-activity patterns on personal exposure to black carbon. *Atmospheric Environment*, 45(21):3594–3602.

Dons, E., Temmerman, P., Van Poppel, M., Bellemans, T., Wets, G., and Int Panis, L. (2013a). Street characteristics and traffic factors determining road users' exposure to black carbon. *Science of the Total Environment*, 447:72–9.

Dons, E., Van Poppel, M., Int Panis, L., De Prins, S., Berghmans, P., Koppen, G., and Matheeussen, C. (2014a). Land use regression models as a tool for short, medium and long term exposure to traffic related air pollution. *Science of the Total Environment*, 476-477:378–386.

Dons, E., Van Poppel, M., Kochan, B., Wets, G., and Int Panis, L. (2013b). Modeling temporal and spatial variability of traffic-related air pollution: Hourly land use regression models for black carbon. *Atmospheric Environment*, 74:237–246.

Dons, E., Van Poppel, M., Kochan, B., Wets, G., and Int Panis, L. (2014b). Implementation and validation of a modeling framework to assess personal exposure to black carbon. *Environment International*, 62:64–71.

dos Santos-Juusela, V., Petäjä, T., Kousa, A., and Hämeri, K. (2013). Spatial - temporal variations of particle number concentrations between a busy street and the urban background. *Atmospheric Environment*, 79:324–333.

Drinovec, L., Močnik, G., Zotter, P., Prévôt, A. S. H., Ruckstuhl, C., Coz, E., Rupakheti, M., Sciare, J., Müller, T., Wiedensohler, A., and Hansen, A. D. A. (2015). The "dual-spot" Aethalometer: an improved measurement of aerosol black carbon with real-time loading compensation. *Atmospheric Measurement Techniques*, 8(5):1965–1979.

EEA (2013a). Air Implementation Pilot. Technical report, European Environment Agency.

EEA (2013b). Status of black carbon monitoring in ambient air in Europe. Technical Report 18, European Environment Agency.

EEA (2015a). *Air quality in Europe - 2015 report (EEA Report Nº 5/2015)*. Number 4. European Environment Agency.

EEA (2015b). SOER 2015 - The European environment - state and outlook 2015. Technical report, European Environment Agency, Copenhagen, Denmark.

Eeftens, M., Beekhuizen, J., Beelen, R., Wang, M., Vermeulen, R., Brunekreef, B., Huss, A., and Hoek, G. (2013). Quantifying urban street configuration for improvements in air pollution models. *Atmospheric Environment*, 72:1–9.

Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Ddel, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Gražulevičien, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Nádor, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.-Y., Yli-Tuomi, T., Varró, M. J., Vienneau, D., von Klot, S., Wolf, K., Brunekreef, B., and Hoek, G. (2012a). Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. *Environmental Science & Technology*, 46(20):11195–205.

Eeftens, M., Tsai, M.-Y., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G., Cirach, M., Cyrys, J., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Ddel, A., Eriksen, K., Galassi, C., Gražulevičien, R., Grivas, G., Heinrich, J., Hoffmann, B., Iakovides, M., Ineichen, A., Katsouyanni, K., Korek, M., Krämer, U., Kuhlbusch, T., Lanki, T., Madsen, C., Meliefste, K., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Oldenwening, M., Pennanen, A., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Udvardy, O., Vaskövi, É., Weinmayr, G., Brunekreef, B., and Hoek, G. (2012b). Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2 - Results of the ESCAPE project. *Atmospheric Environment*, 62:303–317.

Elen, B., Peters, J., Van Poppel, M., Bleux, N., Theunis, J., Reggente, M., and Standaert, A. (2013). The aeroflex: a bicycle for mobile air quality measurements. *Sensors*, 13(1):221–240.

Elen, B., Theunis, J., Ingarra, S., Molino, A., Van den Bossche, J., Reggente, M., and Loreto, V. (2012). The EveryAware SensorBox: a tool for community-based air quality monitoring. In *Workshop Sensing a Changing World*, number 1, pages 1–7.

Elminir, H. K. (2005). Dependence of urban air pollutants on meteorology. *Science of the Total Environment*, 350(1-3):225–237.

EPA (2012). Report to Congress on Black Carbon. Technical report, U.S. Environmental Protection Agency.

EPA (2013). DRAFT Roadmap for Next Generation Air Monitoring. Technical report, U.S. Environmental Protection Agency.

EPA (2016a). Air Sensor Toolbox for Citizen Scientists.

EPA (2016b). National Ambient Air Quality Standards (NAAQS).

EU (2008). Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe.

FPS Economy and Statistics Division (2008). Time Use Survey.

Fuller, G. W., Sciare, J., Lutz, M., Moukhtar, S., and Wagener, S. (2013). New Directions: Time to tackle urban wood burning? *Atmospheric Environment*, 68:295–296.

Gerharz, L. E., Klemm, O., Broich, A. V., and Pebesma, E. (2013). Spatio-temporal modelling of individual exposure to air pollution and its uncertainty. *Atmospheric Environment*, 64:56–65.

Gerharz, L. E., Krüger, A., and Klemm, O. (2009). Applying indoor and outdoor modeling techniques to estimate individual exposure to PM2.5 from personal GPS profiles and diaries: a pilot study. *Science of the Total Environment*, 407(18):5184–93.

Ghassoun, Y., Ruths, M., Löwner, M.-O., and Weber, S. (2015). Intra-urban variation of ultrafine particles as evaluated by process related land use and pollutant driven regression modelling. *Science of the Total Environment*, 536:150–160.

Ghosh, J. K. C., Wilhelm, M., Su, J., Goldberg, D., Cockburn, M., Jerrett, M., and Ritz, B. (2012). Assessing the influence of traffic-related air pollution on risk of term low birth weight on the basis of land-use-based regression models and measures of air toxics. *American Journal of Epidemiology*, 175(12):1262–1274.

Glasgow, M. L., Rudra, C. B., Yoo, E.-H., Demirbas, M., Merriman, J., Nayak, P., Crabtree-Ide, C., Szpiro, A. A., Rudra, A., Wactawski-Wende, J., and Mu, L. (2014). Using smartphones to collect time-activity data for long-term personal-level air pollution exposure assessment. *Journal of exposure science & environmental epidemiology*.

Goel, A. and Kumar, P. (2014). A review of fundamental drivers governing the emissions, dispersion and exposure to vehicle-emitted nanoparticles at signalised traffic intersections. *Atmospheric Environment*, 97:316–331.

Grundström, M., Hak, C., Chen, D., Hallquist, M., and Pleijel, H. (2015). Variation and co-variation of PM10, particle number concentration, NOx and NO2 in the urban air Relationships with wind speed, vertical temperature gradient and weather type. *Atmospheric Environment*, 120:317–327.

Gryparis, A., Dimakopoulou, K., Pedeli, X., and Katsouyanni, K. (2014). Spatio-temporal semiparametric models for NO2 and PM10 concentration levels in Athens, Greece. *Science of the Total Environment*, 479-480C:21–30.

Gu, J., Schnelle-Kreis, J., Pitz, M., Diemer, J., Reller, A., Zimmermann, R., Soentgen, J., Peters, A., and Cyrys, J. (2013). Spatial and temporal variability of PM10 sources in Augsburg, Germany. *Atmospheric Environment*.

Hagemann, R., Corsmeier, U., Kottmeier, C., Rinke, R., Wieser, A., and Vogel, B. (2014). Spatial variability of particle number concentrations and NOx in the Karlsruhe (Germany) area obtained with the mobile laboratory AERO-TRAM'. *Atmospheric Environment*, 94:341–352.

Hagler, G. S. W., Baldauf, R. W., Thoma, E., Long, T., Snow, R., Kinsey, J., Oudejans, L., and Gullett, B. (2009). Ultrafine particles near a major roadway in Raleigh, North Carolina: Downwind attenuation and correlation with traffic-related pollutants. *Atmospheric Environment*, 43(6):1229–1234.

Hagler, G. S. W., Lin, M.-Y., Khlystov, A., Baldauf, R. W., Isakov, V., Faircloth, J., and Jackson, L. E. (2012). Field investigation of roadside vegetative and structural barrier impact on near-road ultrafine particle concentrations under a variety of wind conditions. *Science of the Total Environment*, 419:7–15.

Hagler, G. S. W., Thoma, E. D., and Baldauf, R. W. (2010). High-resolution mobile monitoring of carbon monoxide and ultrafine particle concentrations in a near-road environment. *Journal of the Air & Waste Management Association*, 60(3):328–336.

Hagler, G. S. W., Yelverton, T. L., Vedantham, R., Hansen, A. D., and Turner, J. R. (2011). Post-processing method to reduce noise while preserving high time resolution in Aethalometer real-time black carbon data. *Aerosol and Air Quality Research*, 11(5):539–546.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J. (2009). The formation, properties and impact of secondary organic aerosol: current and emerging issues. *Atmospheric Chemistry and Physics*, 9(14):5155–5236.

Hampel, R., Peters, A., Beelen, R., Brunekreef, B., Cyrys, J., de Faire, U., de Hoogh, K., Fuks, K., Hoffmann, B., Huls, A., Imboden, M., Jedynska, A., Kooter, I., Koenig, W., Kunzli, N., Leander, K., Magnusson, P., Mannisto, S., Penell, J., Pershagen, G. G., Phuleria, H., Probst-Hensch, N., Pundt, N., Schaffner, E., Schikowski, T., Sugiri, D., Tiittanen, P., Tsai, M.-Y., Wang, M., Wolf, K., Lanki, T., Hüls, A., Imboden, M., Jedynska, A., Kooter, I., Koenig, W., Künzli, N., Leander, K., Magnusson, P., Männistö, S., Penell, J., Pershagen, G. G., Phuleria, H., Probst-Hensch, N., Pundt, N., Schaffner, E., Schikowski, T., Sugiri, D., Tiittanen, P., Tsai, M.-Y., Wang, M., Wolf, K., and Lanki, T. (2015). Long-term effects of elemental composition of particulate matter on inflammatory blood markers in European cohorts. *Environment International*, 82:76–84.

Hankey, S. and Marshall, J. D. (2015a). Land Use Regression models of on-road particulate air pollution (Particle Number, Black Carbon, PM2.5, Particle Size) using mobile monitoring. *Environmental Science & Technology*, 49(15):9194–9202.

Hankey, S. and Marshall, J. D. (2015b). On-bicycle exposure to particulate air pollution: Particle number, black carbon, PM2.5, and particle size. *Atmospheric Environment*, 122:65–73.

Hansen, A., Rosen, H., and Novakov, T. (1984). The aethalometer An instrument for the real-time measurement of optical absorption by aerosol particles. *Science of the Total Environment*, 36:191–196.

Harrison, R. M., Brunekreef, B., Keuken, M., Denier van der Gon, H., and Querol, X. (2014). New directions: Cleaning the air: will the European Commission's Clean Air Policy Package of December 2013 deliver? *Atmospheric Environment*.

Hasenfratz, D., Saukh, O., Walser, C., Hueglin, C., Fierz, M., Arn, T., Beutel, J., and Thiele, L. (2015). Deriving high-resolution urban air pollution maps using mobile sensor nodes. *Pervasive and Mobile Computing*, 16:268–285.

Heal, M. and Hammonds, M. (2014). Insights into the composition and sources of rural, urban and roadside carbonaceous PM10. *Environmental Science & Technology*, 48(16):8995–9003.

Heal, M. R., Kumar, P., and Harrison, R. M. (2012). Particles, air quality, policy and health. *Chemical Society reviews*.

HEI (2010). Traffic-related air pollution: A critical review of the literature on emissions, exposure, and health effects. Technical Report January, Health Effects Institute, Boston.

Henderson, S. B., Beckerman, B., Jerrett, M., and Brauer, M. (2007). Application of Land Use Regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology*, 41(7):2422–2428.

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., and Briggs, D. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, 42(33):7561–7578.

Hoek, G., Boogaard, H., Knol, A., de Hartog, J., Slottje, P., Ayres, J. G., Borm, P., Brunekreef, B., Donaldson, K., Forastiere, F., Holgate, S., Kreyling, W. G., Nemery, B., Pekkanen, J., Stone, V., Wichmann, H.-E., and van der Sluijs, J. (2010). Concentration response functions for ultrafine particles and all-cause mortality and hospital admissions: Results of a European expert panel elicitation. *Environmental Science & Technology*, 44(1):476–82.

Hoek, G., Meliefste, K., Cyrys, J., Lewné, M., Bellander, T., Brauer, M., Fischer, P., Gehring, U., Heinrich, J., Van Vliet, P., and Brunekreef, B. (2002). Spatial variability

of fine particle concentrations in three European areas. *Atmospheric Environment*, 36(25):4077–4088.

Holstius, D. M., Pillarisetti, A., Smith, K. R., and Seto, E. (2014). Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California. *Atmospheric Measurement Techniques*, 7(4):1121–1131.

Hsu, H.-H., Adamkiewicz, G., Houseman, E. A., Spengler, J. D., and Levy, J. I. (2014). Using mobile monitoring to characterize roadway and aircraft contributions to ultrafine particle concentrations near a mid-sized airport. *Atmospheric Environment*, 89:688–695.

Hu, S., Paulson, S. E., Fruin, S. A., Kozawa, K., Mara, S., and Winer, A. M. (2012). Observation of elevated air pollutant concentrations in a residential neighborhood of Los Angeles California using a mobile platform. *Atmospheric Environment*, 51:311–319.

Hudda, N., Gould, T., Hartin, K., Larson, T. V., and Fruin, S. A. (2014). Emissions from an international airport increase particle number concentrations 4-fold at 10 km downwind. *Environmental Science & Technology*, 48(12):6628–35.

IARC (2013). Diesel and Gasoline Engine Exhausts and Some Nitroarenes. Technical report, International Agency for Research on Cancer.

IARC (2015). Outdoor Air Pollution. Technical report, International Agency for Research on Cancer.

Int Panis, L., de Geus, B., Vandenbulcke, G., Willems, H., Degraeuwe, B., Bleux, N., Mishra, V., Thomas, I., and Meeusen, R. (2010). Exposure to particulate matter in traffic: A comparison of cyclists and car passengers. *Atmospheric Environment*, 44(19):2263–2270.

IRCEL (2015). JAARRAPPORT Luchtkwaliteit in België 2012. Technical report, Belgian Interregional Environment Agency.

Janssen, N. A., Gerlofs-Nijland, M. E., Lanki, T., Salonen, R. O., Cassee, F., Hoek, G., Fischer, P., Brunekreef, B., and Krzyzanowski, M. (2012). Health effects of black carbon. Technical report, World Health Organization.

Janssen, N. A. H., Hoek, G., Simic-Lawson, M., Fischer, P., van Bree, L., ten Brink, H., Keuken, M., Atkinson, R., Anderson, H., Brunekreef, B., and Cassee, F. R. (2011). Black carbon as an additional indicator of the adverse health effects of airborne particles compared to PM10 and PM2.5. *Environmental Health Perspectives*, 119(12):1691–1699.

Janssen, S., Dumont, G., Fierens, F., and Mensink, C. (2008). Spatial interpolation of air pollution measurements using CORINE land cover data. *Atmospheric Environment*, 42(20):4884–4903.

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., and Giovis, C. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology*, 15(2):185–204.

Ježek, I., Drinovec, L., Ferrero, L., Carriero, M., and Močnik, G. (2015). Determination of car on-road black carbon and particle number emission factors and comparison between mobile and stationary measurements. *Atmospheric Measurement Techniques*, 8(1):43–55.

Jiao, W., Hagler, G., Williams, R., Sharpe, B., Brown, R., Garver, D., Judge, R., Caudill, M., Rickard, J., Davis, M., Weinstock, L., and Buckley, K. (2015). Community Air Sensor Network (CAIRSENSE) project: lower cost, continuous ambient monitoring methods. In *108 th Annual Meeting of the Air & Waste Management Association*.

Johnson, K. K., Bergin, M. H., Russell, A. G., and Hagler, G. S. W. (2016). Using low cost sensors to measure ambient particulate matter concentrations and on-road emissions factors. *Atmospheric Measurement Techniques Discussions*, pages 1–22.

Jones, A. M., Harrison, R. M., and Baker, J. (2010). The wind speed dependence of the concentrations of airborne particulate matter and NOx. *Atmospheric Environment*, 44(13):1682–1690.

Kamboures, M. A., Rieger, P. L., Zhang, S., Sardar, S. B., Chang, M.-C. O., Huang, S.-M., Dzhema, I., Fuentes, M., Benjamin, M. T., Hebert, A., and Ayala, A. (2015). Evaluation of a method for measuring vehicular PM with a composite filter and a real-time BC instrument. *Atmospheric Environment*, 123:63–71.

Kanaroglou, P. S., Adams, M. D., De Luca, P. F., Corr, D., and Sohel, N. (2013). Estimation of sulfur dioxide air pollution concentrations with a spatial autoregressive model. *Atmospheric Environment*, 79:421–427.

Kapadia, A., Kotz, D., and Triandopoulos, N. (2009). Opportunistic sensing: Security challenges for the new paradigm. In *2009 First International Communication Systems and Networks and Workshops*, pages 1–10. IEEE.

Karanasiou, A., Viana, M., Querol, X., Moreno, T., and de Leeuw, F. (2014). Assessment of personal exposure to particulate air pollution during commuting in European cities-Recommendations and policy implications. *Science of the Total Environment*, 490C:785–797.

Kaur, S., Clark, R., Walsh, P., Arnold, S., Colvile, R., and Nieuwenhuijsen, M. (2006). Exposure visualisation of ultrafine particle counts in a transport microenvironment. *Atmospheric Environment*, 40(2):386–398.

Kaur, S., Nieuwenhuijsen, M., and Colvile, R. (2005). Personal exposure of street canyon intersection users to PM2.5, ultrafine particle counts and carbon monoxide in Central London, UK. *Atmospheric Environment*, 39(20):3629–3641.

Kaur, S., Nieuwenhuijsen, M., and Colvile, R. (2007). Fine particulate matter and carbon monoxide exposure concentrations in urban street transport microenvironments. *Atmospheric Environment*, 41(23):4781–4810.

Kingham, S., Longley, I., Salmond, J., Pattinson, W., and Shrestha, K. (2013). Variations in exposure to traffic pollution while travelling by different modes in a low density, less congested city. *Environmental Pollution*, 181:211–218.

Klompmaker, J. O., Montagne, D. R., Meliefste, K., Hoek, G., and Brunekreef, B. (2015). Spatial variation of ultrafine particles and black carbon in two cities: results from a short-term measurement campaign. *Science of the Total Environment*, 508:266–275.

Knibbs, L. D., Cole-Hunter, T., and Morawska, L. (2011). A review of commuter exposure to ultrafine particles and its health effects. *Atmospheric Environment*, 45(16):2611–2622.

Knibbs, L. D., Hewson, M. G., Bechle, M. J., Marshall, J. D., and Barnett, A. G. (2014). A national satellite-based land-use regression model for air pollution exposure assessment in Australia. *Environmental research*, 135C:204–211.

Koehler, K. A. and Peters, T. M. (2015). New methods for personal exposure monitoring for airborne particles. *Current Environmental Health Reports*, 2(4):399–411.

Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260):583–621.

Kuhlbusch, T. A., Quincey, P., Fuller, G. W., Kelly, F., Mudway, I., Viana, M., Querol, X., Alastuey, A., Katsouyanni, K., Weijers, E., Borowiak, A., Gehrig, R., Hueglin, C., Bruckmann, P., Favez, O., Sciare, J., Hoffmann, B., EspenYttri, K., Torseth, K., Sager, U., Asbach, C., and Quass, U. (2014). New Directions: The future of European urban air quality monitoring. *Atmospheric Environment*, 87:258–260.

Kumar, P., Ketzel, M., Vardoulakis, S., Pirjola, L., and Britter, R. (2011). Dynamics and dispersion modelling of nanoparticles from road traffic in the urban atmospheric environment – A review. *Journal of Aerosol Science*, 42(9):580–603.

Kumar, P., Morawska, L., Birmili, W., Paasonen, P., Hu, M., Kulmala, M., Harrison, R. M., Norford, L., and Britter, R. (2014). Ultrafine particles in cities. *Environment International*, 66:1–10.

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., and Britter, R. (2015). The rise of low-cost sensing for managing air pollution in cities. *Environment International*, 75:199–205.

Kumar, P., Robins, A., Vardoulakis, S., and Britter, R. (2010). A review of the characteristics of nanoparticles in the urban atmosphere and the prospects for developing regulatory controls. *Atmospheric Environment*, 44(39):5035–5052.

Larson, T., Henderson, S. B., and Brauer, M. (2009). Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. *Environmental Science & Technology*, 43(13):4672–4678.

Lefebvre, W., Van Poppel, M., Maiheu, B., Janssen, S., and Dons, E. (2013). Evaluation of the RIO-IFDM-street canyon model chain. *Atmospheric Environment*, 77:325–337.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 525(7569):367–371.

Lewandowska, A., Falkowska, L., Murawiec, D., Pryputniewicz, D., Burska, D., and Bełdowska, M. (2010). Elemental and organic carbon in aerosols over urbanized coastal region (southern Baltic Sea, Gdynia). *Science of the Total Environment*, 408(20):4761–9.

Li, B., Lei, X.-N., Xiu, G.-L., Gao, C.-Y., Gao, S., and Qian, N.-S. (2015). Personal exposure to black carbon during commuting in peak and off-peak hours in Shanghai. *Science of the Total Environment*, 524-525C:237–245.

Li, L., Wu, J., Ghosh, J. K., and Ritz, B. (2013). Estimating spatiotemporal variability of ambient air pollutant concentrations with a hierarchical model. *Atmospheric Environment*, 71:54–63.

Li, L., Wu, J., Wilhelm, M., and Ritz, B. (2012). Use of generalized additive models and cokriging of spatial residuals to improve land-use regression estimates of nitrogen oxides in Southern California. *Atmospheric Environment*, 55:220–228.

Liang, M. S., Keener, T. C., Birch, M. E., Baldauf, R., Neal, J., and Yang, Y. J. (2013). Low-wind and other microclimatic factors in near-road black carbon variability: A case study and assessment implications. *Atmospheric Environment*, 80:204–215.

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., Blore, J. D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M., Brauer, M., Brooks, P., Bruce, N. G., Brunekreef, B., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Bull, F., Burnett, R. T., Byers, T. E., Calabria, B., Carapetis, J., Carnahan, E., Chafe, Z., Charlson, F., Chen, H., Chen, J. S., Cheng, A. T.-A., Child, J. C., Cohen, A., Colson, K. E., Cowie, B. C., Darby, S., Darling, S., Davis, A., Degenhardt, L., Dentener, F., Des Jarlais, D. C., Devries, K., Dherani, M., Ding, E. L., Dorsey, E. R., Driscoll, T., Edmond, K., Ali, S. E., Engell, R. E., Erwin, P. J., Fahimi, S., Falder, G., Farzadfar, F., Ferrari, A., Finucane, M. M., Flaxman, S., Fowkes, F. G. R., Freedman, G., Freeman, M. K., Gakidou, E., Ghosh, S., Giovannucci, E., Gmel, G., Graham, K., Grainger, R., Grant, B., Gunnell, D., Gutierrez, H. R., Hall, W., Hoek, H. W., Hogan, A., Hosgood, H. D., Hoy, D., Hu, H., Hubbell, B. J., Hutchings, S. J., Ibeanusi, S. E., Jacklyn, G. L., Jasrasaria, R., Jonas, J. B., Kan, H., Kanis, J. A., Kassebaum, N.,

Kawakami, N., Khang, Y.-H., Khatibzadeh, S., Khoo, J.-P., Kok, C., Laden, F., Lalloo, R., Lan, Q., Lathlean, T., Leasher, J. L., Leigh, J., Li, Y., Lin, J. K., Lipshultz, S. E., London, S., Lozano, R., Lu, Y., Mak, J., Malekzadeh, R., Mallinger, L., Marcenes, W., March, L., Marks, R., Martin, R., McGale, P., McGrath, J., Mehta, S., Mensah, G. A., Merriman, T. R., Micha, R., Michaud, C., Mishra, V., Mohd Hanafiah, K., Mokdad, A. A., Morawska, L., Mozaffarian, D., Murphy, T., Naghavi, M., Neal, B., Nelson, P. K., Nolla, J. M., Norman, R., Olives, C., Omer, S. B., Orchard, J., Osborne, R., Ostro, B., Page, A., Pandey, K. D., Parry, C. D. H., Passmore, E., Patra, J., Pearce, N., Pelizzari, P. M., Petzold, M., Phillips, M. R., Pope, D., Pope, C. A., Powles, J., Rao, M., Razavi, H., Rehfuess, E. A., Rehm, J. T., Ritz, B., Rivara, F. P., Roberts, T., Robinson, C., Rodriguez-Portales, J. A., Romieu, I., Room, R., Rosenfeld, L. C., Roy, A., Rushton, L., Salomon, J. A., Sampson, U., Sanchez-Riera, L., Sanman, E., Sapkota, A., Seedat, S., Shi, P., Shield, K., Shivakoti, R., Singh, G. M., Sleet, D. A., Smith, E., Smith, K. R., Stapelberg, N. J. C., Steenland, K., Stöckl, H., Stovner, L. J., Straif, K., Straney, L., Thurston, G. D., Tran, J. H., Van Dingenen, R., van Donkelaar, A., Veerman, J. L., Vijayakumar, L., Weintraub, R., Weissman, M. M., White, R. A., Whiteford, H., Wiersma, S. T., Wilkinson, J. D., Williams, H. C., Williams, W., Wilson, N., Woolf, A. D., Yip, P., Zielinski, J. M., Lopez, A. D., Murray, C. J. L., Ezzati, M., AlMazroa, M. A., and Memish, Z. A. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 380(9859):2224–2260.

Louwies, T., Nawrot, T., Cox, B., Dons, E., Penders, J., Provost, E., Panis, L. I., and De Boever, P. (2015). Blood pressure changes in association with black carbon exposure in a panel of healthy adults are independent of retinal microcirculation. *Environment international*, 75:81–6.

MacNaughton, P., Melly, S., Vallarino, J., Adamkiewicz, G., and Spengler, J. D. (2014). Impact of bicycle route type on exposure to traffic-related air pollution. *Science of the Total Environment*, 490:37–43.

Maenhout, W. and Cafmeyer, J. (1998). Long-term atmospheric aerosol study at urban and rural sites in Belgium using multi-elemental analysis by particle-induced X-ray emission spectrometry and short-irradiation instrumental neutron activation analysis. *X-ray Spectrometry*, 27(4):236–246.

Maynard, D., Coull, B. A., Gryparis, A., and Schwartz, J. (2007). Mortality risk associated with short-term exposure to traffic particles and sulfates. *Environmental Health Perspectives*, 115(5):751–5.

McCreanor, J., Cullinan, P., Nieuwenhuijsen, M. J., Stewart-Evans, J., Malliarou, E., Jarup, L., Harrington, R., Svartengren, M., Han, I.-K., Ohman-Strickland, P., Chung, K. F., and Zhang, J. (2007). Respiratory effects of exposure to diesel traffic in persons with asthma. *The New England Journal of Medicine*, 357(23):2348–2358.

McNabola, A., Broderick, B., and Gill, L. (2008). Relative exposure to fine particulate matter and VOCs between transport microenvironments in Dublin: Personal exposure and uptake. *Atmospheric Environment*, 42(26):6496–6512.

Mead, M., Popoola, O., Stewart, G., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J., McLeod, M., Hodgson, T., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J., and Jones, R. (2013). The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment*, 70:186–203.

Merbitz, H., Fritz, S., and Schneider, C. (2012). Mobile measurements and regression modeling of the spatial particulate matter variability in an urban area. *Science of the Total Environment*, 438:389–403.

Mishra, V. K., Kumar, P., Van Poppel, M., Bleux, N., Frijns, E., Reggente, M., Berghmans, P., Int Panis, L., and Samson, R. (2012). Wintertime spatio-temporal variation of ultrafine particles in a Belgian city. *Science of the Total Environment*, 431:307–13.

Mölter, A., Lindley, S., de Vocht, F., Agius, R., Kerry, G., Johnson, K., Ashmore, M., Terry, A., Dimitroulopoulou, S., and Simpson, A. (2012). Performance of a microenviromental model for estimating personal NO2 exposure in children. *Atmospheric Environment*, 51(2):225–233.

Montagne, D., Hoek, G., Nieuwenhuijsen, M., Lanki, T., Pennanen, A., Portella, M., Meliefste, K., Eeftens, M., Yli-Tuomi, T., Cirach, M., and Brunekreef, B. (2013). Agreement of land use regression models with personal exposure measurements of particulate matter and nitrogen oxides air pollution. *Environmental Science & Technology*, 47(15):8523–31.

Montagne, D., Hoek, G., Nieuwenhuijsen, M., Lanki, T., Pennanen, A., Portella, M., Meliefste, K., Wang, M., Eeftens, M., Yli-Tuomi, T., Cirach, M., and Brunekreef, B. (2014). The association of LUR modeled PM2.5 elemental composition with personal exposure. *Science of the Total Environment*, 493:298–306.

Montagne, D. R., Hoek, G., Klompmaker, J. O., Wang, M., Meliefste, K., and Brunekreef, B. (2015). Land use regression models for ultrafine particles and black carbon based on short-term monitoring predict past spatial variation. *Environmental Science & Technology*, 49(14):8712–8720.

Morawska, L., Afshari, A., Bae, G. N., Buonanno, G., Chao, C. Y. H., Hänninen, O., Hofmann, W., Isaxon, C., Jayaratne, E. R., Pasanen, P., Salthammer, T., Waring, M., and Wierzbicka, A. (2013). Indoor aerosols: from personal exposure to risk assessment. *Indoor air*, 23(6):462–87.

Morawska, L., Ristovski, Z., Jayaratne, E., Keogh, D., and Ling, X. (2008). Ambient nano and ultrafine particles from motor vehicle emissions: Characteristics, ambient processing and implications on human exposure. *Atmospheric Environment*, 42(35):8113–8138.

Mueller, M., Hasenfratz, D., Saukh, O., Fierz, M., and Hueglin, C. (2016). Statistical modelling of particle number concentration in Zurich at high spatio-temporal resolution utilizing data from a mobile sensor network. *Atmospheric Environment*, 126:171–181.

Nieuwenhuijsen, M. J., Donaire-Gonzalez, D., Rivas, I., de Castro, M., Cirach, M., Hoek, G., Seto, E., Jerrett, M., and Sunyer, J. (2015). Variability in and agreement between modeled and personal continuously measured black carbon levels using novel smartphone and sensor technologies. *Environmental Science & Technology*, 49(5):2977–2982.

Nikolova, I., Janssen, S., Vrancken, K., Vos, P., Mishra, V., and Berghmans, P. (2011). Size resolved ultrafine particles emission model–A continues size distribution approach. *Science of the Total Environment*, 409(18):3492–3499.

Ning, Z., Chan, K., Wong, K., Westerdahl, D., Močnik, G., Zhou, J., and Cheung, C. (2013). Black carbon mass size distributions of diesel exhaust and urban aerosols measured using differential mobility analyzer in tandem with Aethalometer. *Atmospheric Environment*, 80:31–40.

Padró-Martínez, L. T., Patton, A. P., Trull, J. B., Zamore, W., Brugge, D., and Durant, J. L. (2012). Mobile monitoring of particle number concentration and other traffic-related air pollutants in a near-highway neighborhood over the course of a year. *Atmospheric Environment*, 61(null):253–264.

Panis, L. I. (2010). New Directions: Air pollution epidemiology can benefit from activity-based models. *Atmospheric Environment*, 44(7):1003–1004.

Patel, M. M. M., Chillrud, S. N. S., Correa, J. C. J., Hazi, Y., Feinberg, M., KC, D., Prakash, S., Ross, J. M., Levy, D., and Kinney, P. P. L. (2010). Traffic-related particulate matter and acute respiratory symptoms among New York City area adolescents. *Environmental Health Perspectives*, 118(9):1338–43.

Pattinson, W., Longley, I., and Kingham, S. (2014). Using mobile monitoring to visualise diurnal variation of traffic pollutants across two near-highway neighbourhoods. *Atmospheric Environment*, 94:782–792.

Patton, A. P., Collins, C., Naumova, E. N., Zamore, W., Brugge, D., and Durant, J. L. (2014a). An hourly regression model for ultrafine particles in a near-highway urban area. *Environmental Science & Technology*, (48):3272–3280.

Patton, A. P., Perkins, J., Zamore, W., Levy, J. I., Brugge, D., and Durant, J. L. (2014b). Spatial and temporal differences in traffic-related air pollution in three urban neighborhoods near an interstate highway. *Atmospheric Environment*, 99:309–321.

Patton, A. P., Zamore, W., Naumova, E. N., Levy, J. I., Brugge, D., and Durant, J. L. (2015). Transferability and generalizability of regression models of ultrafine particles in urban neighborhoods in the Boston Aarea. *Environmental Science & Technology*, 49(10):6051–6060.

Pearce, J. L., Beringer, J., Nicholls, N., Hyndman, R. J., and Tapper, N. J. (2011). Quantifying the influence of local meteorology on air quality using generalized additive models. *Atmospheric Environment*, 45(6):1328–1336.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Peters, J., Sîrbu, A., Van den Bossche, J., Ingarra, S., De Baets, B., and Theunis, J. (2013a). Development of a low-cost mobile sensor-system for participatory measurements of urban air quality. In *COST Action TD1105 EuNetAir 1st International Workshop, Open Satellite Workshop inside Transducers 2013, Barcelona, 20 June 2013*, Barcelona.

Peters, J., Theunis, J., Van Poppel, M., and Berghmans, P. (2013b). Monitoring PM10 and ultrafine particles in urban environments using mobile measurements. *Aerosol and Air Quality Research*, 13:509–522.

Peters, J., Van den Bossche, J., Reggente, M., Van Poppel, M., De Baets, B., and Theunis, J. (2014). Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. *Atmospheric Environment*, 92:31–43.

Petzold, A., Ogren, J. A., Fiebig, M., Laj, P., Li, S.-M., Baltensperger, U., Holzer-Popp, T., Kinne, S., Pappalardo, G., Sugimoto, N., Wehrli, C., Wiedensohler, A., and Zhang, X.-Y. (2013). Recommendations for reporting "black carbon" measurements. *Atmospheric Chemistry and Physics*, 13(16):8365–8379.

Petzold, A., Schloesser, H., Sheridan, P. J., Arnott, W. P., Ogren, J. A., and Virkkula, A. (2005). Evaluation of multiangle absorption photometry for measuring aerosol light absorption. *Aerosol Science and Technology*, 39(1):40–51.

Petzold, A. and Schonlinner, M. (2004). Multi-angle absorption photometry - a new method for the measurement of aerosol light absorption and atmospheric black carbon. *Journal of Aerosol Science*, 35(4):421–441.

Physick, W., Powell, J., Cope, M., Boast, K., and Lee, S. (2011). Measurements of personal exposure to NO2 and modelling using ambient concentrations and activity data. *Atmospheric Environment*, 45(12):2095–2102.

Piedrahita, R., Xiang, Y., Masson, N., Ortega, J., Collier, A., Jiang, Y., Li, K., Dick, R. P., Lv, Q., Hannigan, M., and Shang, L. (2014). The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques*, 7(10):3325–3336.

Pirjola, L., Pajunoja, A., Walden, J., Jalkanen, J.-P., Rönkkö, T., Kousa, A., and Koskentalo, T. (2014). Mobile measurements of ship emissions in two harbour areas in Finland. *Atmospheric Measurement Techniques*, 7:149–161.

Pope, C. A., Ezzati, M., and Dockery, D. W. (2009). Fine-particulate air pollution and life expectancy in the United States. *The New England Journal of Medicine*, 360(4):376–86.

Provost, E. B., Louwies, T., Cox, B., Op 't Roodt, J., Solmi, F., Dons, E., Int Panis, L., De Boever, P., and Nawrot, T. S. (2016). Short-term fluctuations in personal black carbon exposure are associated with rapid changes in carotid arterial stiffening. *Environment International*, 88:228–234.

PyData Development Team (2014). pandas: Python Data Analysis Library.

QGIS Development Team (2014). *QGIS Geographic Information System*. Open Source Geospatial Foundation.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raaschou-Nielsen, O., Andersen, Z. J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M. J., Brunekreef, B., Xun, W. W., Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., Modig, L., Oudin, A., Oftedal, B., Schwarze, P. E., Nafstad, P., De Faire, U., Pedersen, N. L., Ostenson, C.-G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G. G., Eriksen, K. T., Sørensen, M., Tjønneland, A., Ellermann, T., Eeftens, M., Peeters, P. H., Meliefste, K., Wang, M., Bueno-de Mesquita, B., Key, T. J., de Hoogh, K., Concin, H., Nagel, G., Vilier, A., Grioni, S., Krogh, V., Tsai, M.-Y. Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Trichopoulou, A., Bamia, C., Vineis, P., and Hoek, G. (2013). Air pollution and lung cancer incidence in 17 European cohorts: Prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *The Lancet Oncology*, 14(9):813–822.

Ragettli, M. S., Ducret-Stich, R. E., Foraster, M., Morelli, X., Aguilera, I., Basagaña, X., Corradi, E., Ineichen, A., Tsai, M.-Y., Probst-Hensch, N., Rivera, M., Slama, R., Künzli, N., and Phuleria, H. C. (2014). Spatio-temporal variation of urban ultrafine particle number concentrations. *Atmospheric Environment*, 96:275–283.

Ramsey, N. R., Klein, P. M., and Moore, B. (2014). The impact of meteorological parameters on urban air quality. *Atmospheric Environment*, 86:58–67.

Reche, C., Querol, X., Alastuey, A., Viana, M., Pey, J., Moreno, T., Rodríguez, S., González, Y., Fernández-Camacho, R., de la Rosa, J., Dall'Osto, M., Prévôt, A. S. H., Hueglin, C., Harrison, R. M., and Quincey, P. (2011). New considerations for PM, Black Carbon and particle number concentration for air quality monitoring across different European cities. *Atmospheric Chemistry and Physics*, 11(13):6207–6227.

Rose, N., Cowie, C., Gillett, R., and Marks, G. B. (2011). Validation of a spatiotemporal land use regression model incorporating fixed site monitors. *Environmental Science & Technology*, 45(1):294–9.

Ruths, M., von Bismarck-Osten, C., and Weber, S. (2014). Measuring and modelling the local-scale spatio-temporal variation of urban particle number size distributions and black carbon. *Atmospheric Environment*, 96:37–49.

Ryan, P. H., Son, S. Y., Wolfe, C., Lockey, J., Brokamp, C., and LeMasters, G. (2015). A field application of a personal sensor for ultrafine particle exposure in children. *Science of the Total Environment*, 508:366–373.

Sabaliauskas, K., Jeong, C.-H., Yao, X., Reali, C., Sun, T., and Evans, G. J. (2015). Development of a land-use regression model for ultrafine particles in Toronto, Canada. *Atmospheric Environment*, 110:84–92.

Santiago, J. L., Martín, F., and Martilli, A. (2013). A computational fluid dynamic modelling approach to assess the representativeness of urban monitoring stations. *Science of the Total Environment*, 454-455:61–72.

Saraswat, A., Apte, J. S., Kandlikar, M., Brauer, M., Henderson, S. B., Marshall, J. D., Vt, B. C., Apte, J. S., Brauer, M., Henderson, S. B., and Marshall, J. D. (2013). Spatiotemporal land use regression models of fine, ultrafine, and black carbon particulate matter in New Delhi, India. *Environmental Science & Technology*, 47(22):12903–12911.

Schneider, P., Castell, N., Van den Bossche, J., and Lahoz, W. (2015). High-resolution mapping of urban air quality using low-cost sensors. In *COST Action TD1105 - New Sensing Technologies for Air-Pollution Control and Environmental Sustainability, Sofia, 16-18 December 2015*.

SCU-UWE (2013). Science for Environment Policy In-depth Report: Environmental Citizen Science. Technical Report 9, Science Communication Unit, University of the West of England, Bristol. Report produced for the European Commission DG Environment.

Seabold, J. and Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. In van der Walt, S. and Millman, J., editors, *Proceedings of the 9th Python in Science Conference*, pages 57–61.

Seinfeld, J. H. and Pandis, S. N. (2006). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons.

Setton, E., Marshall, J. D., Brauer, M., Lundquist, K. R., Hystad, P., Keller, P., and Cloutier-Fisher, D. (2011). The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. *Journal of Exposure Science & Environmental Epidemiology*, 21(1):42–48.

SGS (2010). Strategische geluidsbelastingskaarten agglomeratie Antwerpen. Rapport in opdracht van Stad Antwerpen, 090357-2-v1, SGS Belgium NV. Technical report.

Sîrbu, A., Becker, M., Caminiti, S., De Baets, B., Elen, B., Francis, L., Gravino, P., Hotho, A., Ingarra, S., Loreto, V., Molino, A., Mueller, J., Peters, J., Ricchiuti, F., Saracino, F., Servedio, V. D. P., Stumme, G., Theunis, J., Tria, F., and Van den

Bossche, J. (2015). Participatory patterns in an international air quality monitoring initiative. *PLOS ONE*, 10(8):e0136763.

Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222.

Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., Shelow, D., Hindin, D. A., Kilaru, V. J., and Preuss, P. W. (2013). The changing paradigm of air pollution monitoring. *Environmental Science & Technology*, 47(20):11369–11377.

Song, S., Wu, Y., Xu, J., Ohara, T., Hasegawa, S., Li, J., Yang, L., and Hao, J. (2013). Black carbon at a roadside site in Beijing: Temporal variations and relationships with carbon monoxide and particle number size distribution. *Atmospheric Environment*, 77:213–221.

Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M., and Bonavitacola, F. (2015). Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide. *Sensors and Actuators B: Chemical*, 215:249–257.

Steinle, S., Reis, S., and Sabel, C. E. (2013). Quantifying human exposure to air pollution-Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Science of the Total Environment*, 443:184–93.

Steinle, S., Reis, S., Sabel, C. E., Semple, S., Twigg, M. M., Braban, C. F., Leeson, S. R., Heal, M. R., Harrison, D., Lin, C., and Wu, H. (2015). Personal exposure monitoring of PM2.5 in indoor and outdoor microenvironments. *Science of the Total Environment*, 508:383–94.

Su, J. G., Hopke, P. K., Tian, Y., Baldwin, N., Thurston, S. W., Evans, K., and Rich, D. Q. (2015a). Modeling particulate matter concentrations measured through mobile monitoring in a deletion/substitution/addition approach. *Atmospheric Environment*, 122:477–483.

Su, J. G., Jerrett, M., Meng, Y.-Y., Pickett, M., and Ritz, B. (2015b). Integrating smart-phone based momentary location tracking with fixed site air quality monitoring for personal exposure assessment. *Science of the Total Environment*, 506-507:518–26.

Sullivan, R. and Pryor, S. (2014). Quantifying spatiotemporal variability of fine particles in an urban environment using combined fixed and mobile measurements. *Atmospheric Environment*, 89:664–671.

Sun, Y., Wang, Z., Fu, P., Jiang, Q., Yang, T., Li, J., and Ge, X. (2013). The impact of relative humidity on aerosol composition and evolution processes during wintertime in Beijing, China. *Atmospheric Environment*, 77:927–934.

Szpiro, A. A., Sampson, P. D., Sheppard, L., Lumley, T., Adar, S. D., and Kaufman, J. D. (2010). Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. *Environmetrics*, 21(6):606–631.

Tan, Y., Robinson, A. L., and Presto, A. A. (2014). Quantifying uncertainties in pollutant mapping studies using the Monte Carlo method. *Atmospheric Environment*, 99:333–340.

Tang, N. W., Apte, J. S., Martien, P. T., and Kirchstetter, T. W. (2015). Measurement of black carbon emissions from in-use diesel-electric passenger locomotives in California. *Atmospheric Environment*, 115:295–303.

Theunis, J., Peters, J., and Elen, B. (2016a). Participatory air quality monitoring in urban environments - reconciling technological challenges and participation. In Loreto, V., Haklay, M., Hotho, A., Servedio, V., Stumme, G., Theunis, J., and Tria, F., editors, *Social Sensing, Opinions and Collective Awareness*. Springer.

Theunis, J., Stevens, M., and Botteldooren, D. (2016b). Sensing the environment. In Loreto, V., Haklay, M., Hotho, A., Servedio, V., Stumme, G., Theunis, J., and Tria, F., editors, *Social Sensing, Opinions and Collective Awareness*. Springer.

Theunis, J., Strous, C., and Shandwick, W. (2015). Gezondheid en schone lucht centraal bij European Respiratory Society en Longfonds. *Tijdschrift LUCHT*, (September, 4):20–23.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Trompetter, W., Grange, S., Davy, P., and Ancelet, T. (2013). Vertical and temporal variations of black carbon in New Zealand urban areas during winter. *Atmospheric Environment*, 75:179–187.

Tseng, Y.-C., Wu, F.-J., and Lai, W.-T. (2013). Opportunistic data collection for disconnected wireless sensor networks by mobile mules. *Ad Hoc Networks*, 11(3):1150–1164.

Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J., and De Baets, B. (2015). Mobile monitoring for mapping spatial variation in urban air quality: development and validation of a methodology based on an extensive dataset. *Atmospheric Environment*, 105:148–161.

Van den Bossche, J., Theunis, J., Elen, B., Peters, J., Botteldooren, D., and De Baets, B. (2016). Opportunistic mobile air pollution monitoring: A case study with city wardens in Antwerp. *Atmospheric Environment*, 141:408–421.

Van Poppel, M., Peters, J., and Bleux, N. (2013). Methodology for setup and data processing of mobile air quality measurements to assess the spatial variability of concentrations in urban environments. *Environmental Pollution*, 183:224–233.

Vardoulakis, S., Fisher, B. E., Pericleous, K., and Gonzalez-Flesca, N. (2003). Modelling air quality in street canyons: a review. *Atmospheric Environment*, 37(2):155–182.

Vardoulakis, S., Gonzalezflesca, N., Fisher, B., and Pericleous, K. (2005). Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris. *Atmospheric Environment*, 39(15):2725–2736.

Vardoulakis, S., Solazzo, E., and Lumbreras, J. (2011). Intra-urban and street scale variability of BTEX, NO2 and O3 in Birmingham, UK: Implications for exposure assessment. *Atmospheric Environment*, 45(29):5069–5078.

Viana, M., Pey, J., Querol, X., Alastuey, A., de Leeuw, F., and Lükewille, A. (2014). Natural sources of atmospheric aerosols influencing air quality across Europe. *Science of the Total Environment*, 472:825–833.

Viana, M., Rivas, I., Reche, C., Fonseca, A., Pérez, N., Querol, X., Alastuey, A., Álvarez-Pedrerol, M., and Sunyer, J. (2015). Field comparison of portable and stationary instruments for outdoor urban air exposure assessments. *Atmospheric Environment*, 123:220–228.

Viidanoja, J., Sillanpää, M., Laakia, J., Kerminen, V.-M., Hillamo, R., Aarnio, P., and Koskentalo, T. (2002). Organic and black carbon in PM2.5 and PM10: 1 year of data from an urban site in Helsinki, Finland.

Vilcassim, M. J. R., Thurston, G. D., Peltier, R. E., and Gordon, T. (2014). Black carbon and particulate matter (PM 2.5) concentrations in New York City's subway stations. *Environmental Science & Technology*, 48(24):14738–14745.

Virkkula, A., Chi, X., Ding, A., Shen, Y., Nie, W., Qi, X., Zheng, L., Huang, X., Xie, Y., Wang, J., Petäjä, T., and Kulmala, M. (2015). On the interpretation of the loading correction of the aethalometer. *Atmospheric Measurement Techniques*, 8(10):4415–4427.

Virkkula, A., Makela, T., Hillamo, R., Yli-Tuomi, T., Hirsikko, A., Hameri, K., and Koponen, I. (2007). A simple procedure for correcting loading effects of aethalometer data. *Journal of the Air & Waste Management Association*, 57(10):1214–1222.

VMM (2013). CHEMKAR PM10, Chemische karakterisering van fijn stof in Vlaanderen, 2011-2012. Technical report, Vlaamse Milieu Maatschappij.

VMM (2015). Luchtkwaliteit in het Vlaamse Gewest. Jaarverslag Immissiemeetnetten 2014. Technical report, Vlaamse Milieumaatschappij.

von Bismarck-Osten, C., Birmili, W., Ketzel, M., Massling, A., Petäjä, T., and Weber, S. (2013). Characterization of parameters influencing the spatio-temporal variability of urban particle number size distributions in four European cities. *Atmospheric Environment*, 77:415–429.

Vos, P. E., Maiheu, B., Vankerkom, J., and Janssen, S. (2013). Improving local air quality in cities: To tree or not to tree? *Environmental Pollution*, 183:113–122.

Vranckx, S., Vos, P., Maiheu, B., and Janssen, S. (2015). Impact of trees on pollutant dispersion in street canyons: A numerical study of the annual average effects in Antwerp, Belgium. *Science of the Total Environment*, 532:474–83.

Wallace, J., Corr, D., Deluca, P., Kanaroglou, P., and McCarry, B. (2009). Mobile monitoring of air pollution in cities: the case of Hamilton, Ontario, Canada. *Journal of Environmental Monitoring*, 11(5):998–1003.

Wang, M., Beelen, R., Basagana, X., Becker, T., Cesaroni, G., de Hoogh, K., Dedele, A., Declercq, C., Dimakopoulou, K., Eeftens, M., Forastiere, F., Galassi, C., Gražulevičien, R., Hoffmann, B., Heinrich, J., Iakovides, M., Künzli, N., Korek, M., Lindley, S., Mölter, A., Mosler, G., Madsen, C., Nieuwenhuijsen, M., Phuleria, H., Pedeli, X., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Stempfelet, M., Tsai, M.-Y., Lanki, T., Udvardy, O., Varró, M. J., Wolf, K., Weinmayr, G., Yli-Tuomi, T., Hoek, G., and Brunekreef, B. (2013). Evaluation of land use regression models for NO2 and particulate matter in 20 European study areas: the ESCAPE project. *Environmental Science & Technology*, 47(9):4357–64.

Wang, M., Beelen, R., Bellander, T., Birk, M., Cesaroni, G., Cirach, M., Cyrys, J., de Hoogh, K., Declercq, C., Dimakopoulou, K., Eeftens, M., Eriksen, K. T., Forastiere, F., Galassi, C., Grivas, G., Heinrich, J., Hoffmann, B., Ineichen, A., Korek, M., Lanki, T., Lindley, S., Modig, L., Mölter, A., Nafstad, P., Nieuwenhuijsen, M. J., Nystad, W., Olsson, D., Raaschou-Nielsen, O., Ragettli, M., Ranzi, A., Stempfelet, M., Sugiri, D., Tsai, M.-Y., Udvardy, O., Varró, M. J., Vienneau, D., Weinmayr, G., Wolf, K., Yli-Tuomi, T., Hoek, G., and Brunekreef, B. (2014). Performance of multi-city land use regression models for nitrogen dioxide and fine particles. *Environmental Health Perspectives*, 122(8):843–9.

Wang, M., Beelen, R., Eeftens, M., Meliefste, K., Hoek, G., and Brunekreef, B. (2012). Systematic evaluation of land use regression models for NO2. *Environmental Science & Technology*, 46(8):4481–9.

Wang, M., Gehring, U., Hoek, G., Keuken, M., Jonkers, S., Beelen, R., Eeftens, M., Postma, D. S., and Brunekreef, B. (2015). Air pollution and lung function in Dutch children: A comparison of exposure estimates and associations based on Land Use Regression and dispersion exposure modeling approaches. *Environmental Health Perspectives*, 123(2):847–851.

Weber, S. (2009). Spatio-temporal covariation of urban particle number concentration and ambient noise. *Atmospheric Environment*, 43(34):5518–5525.

Weber, S., Kordowski, K., and Kuttler, W. (2013). Variability of particle number concentration and particle size dynamics in an urban street canyon under different meteorological conditions. *Science of the Total Environment*, 449:102–114.

Weichenthal, S., Dufresne, A., Infante-Rivard, C., and Joseph, L. (2008). Determinants of ultrafine particle exposures in transportation environments: findings of an 8-month

survey conducted in Montréal, Canada. *Journal of Exposure Science & Environmental Epidemiology*, 18(6):551–63.

Weichenthal, S., Kulka, R., Dubeau, A., Martin, C., Wang, D., and Dales, R. (2011). Traffic-related air pollution and acute changes in heart rate variability and respiratory function in urban cyclists. *Environmental Health Perspectives*, 119(10):1373–1378.

Weichenthal, S., Ryswyk, K. V., Goldstein, A., Bagg, S., Shekkarizfard, M., and Hatzopoulou, M. (2016a). A land use regression model for ambient ultrafine particles in Montreal, Canada: A comparison of linear regression and a machine learning approach. *Environmental Research*, 146:65–72.

Weichenthal, S., Van Ryswyk, K., Goldstein, A., Shekarrizfard, M., and Hatzopoulou, M. (2016b). Characterizing the spatial distribution of ambient ultrafine particles in Toronto, Canada: A land use regression model. *Environmental Pollution*, 208:241–248.

Weijers, E., Khlystov, A., Kos, G., and Erisman, J. (2004). Variability of particulate matter concentrations along roads and motorways determined by a moving measurement unit. *Atmospheric Environment*, 38(19):2993–3002.

Westerdahl, D., Fruin, S. A., Sax, T., Fine, P. M., and Sioutas, C. (2005). Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los Angeles. *Atmospheric Environment*, 39(20):3597–3610.

WHO (1999). Monitoring ambient air quality for health impact assessment. Technical report, World Health Organization.

WHO (2005). Principles of characterizing and applying human exposure models. Technical Report 3, World Health Organization.

WHO (2006). *Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*. World Health Organization.

WHO (2010). WHO guidelines for indoor air quality: selected pollutants. Technical report, World Health Organization, Regional Office for Europe.

WHO (2013). Review of evidence on health aspects of air pollution REVIHAAP project: final technical report. Technical report, World Health Organization - Regional Office for Europe, Copenhagen.

WHO (2014). Burden of disease from Ambient Air Pollution for 2012. Technical report, World Health Organization.

Williams, R., Kaufman, A., Hanley, T., Rice, J., and Garvey, S. (2014a). Evaluation of field-deployed low cost PM sensors. Technical Report December, U.S. Environmental Protection Agency, Washington, DC.

Williams, R., Long, R., Beaver, M., Kaufman, A., Zeiger, F., Heimbinder, M., Hang, I., Yap, R., Acharya, B., Ginwald, B., Kupcho, K., Robinson, S., Zaouak, O., Aubert, B.,

Hannigan, M., Piedrahita, R., Masson, N., Moran, B., Rook, M., Heppner, P., Cogar, C., Nikzad, N., and Griswold, W. (2014b). Sensor Evaluation Report. Technical report, U.S. Environmental Protection Agency, Washington, DC.

Wilson, J. G., Kingham, S., Pearce, J., and Sturman, A. P. (2005). A review of intraurban variations in particulate air pollution: Implications for epidemiological research. *Atmospheric Environment*, 39(34):6444–6462.

Wu, H., Reis, S., Lin, C., Beverland, I. J., and Heal, M. R. (2015). Identifying drivers for the intra-urban spatial variability of airborne particulate matter components and their interrelationships. *Atmospheric Environment*, 112:306–316.

Wu, J., Wilhelm, M., Chung, J., and Ritz, B. (2011). Comparing exposure assessment methods for traffic-related air pollution in an adverse pregnancy outcome study. *Environmental Research*, 111(5):685–692.

Xu, X., Brook, J., and Guo, Y. (2007). A statistical assessment of saturation and mobile sampling strategies to estimate long-term average concentrations across urban areas. *Journal of the Air & Waste Management Association*, 57(11):1396–1406.

Zhu, Y., Fung, D. C., Kennedy, N., Hinds, W. C., and Eiguren-Fernandez, A. (2008). Measurements of ultrafine particles and other vehicular pollutants inside a mobile exposure system on Los Angeles freeways. *Journal of the Air & Waste Management Association*, 58(3):424–434.

Zuurbier, M., Hoek, G., Oldenwening, M., Lenters, V., Meliefste, K., van den Hazel, P., and Brunekreef, B. (2010). Commuters' exposure to particulate matter air pollution is affected by mode of transport, fuel type, and route. *Environmental Health Perspectives*, 118(6):783–789.

Zwack, L. M., Hanna, S. R., Spengler, J. D., and Levy, J. I. (2011a). Using advanced dispersion models and mobile monitoring to characterize spatial patterns of ultrafine particles in an urban area. *Atmospheric Environment*, 45(28):4822–4829.

Zwack, L. M., Paciorek, C. J., Spengler, J. D., and Levy, J. I. (2011b). Characterizing local traffic contributions to particulate air pollution in street canyons using mobile monitoring techniques. *Atmospheric Environment*, 45(15):2507–2514.

Zwack, L. M., Paciorek, C. J., Spengler, J. D., and Levy, J. I. (2011c). Modeling spatial patterns of traffic-related air pollutants in complex urban terrain. *Environmental Health Perspectives*, 119(6):852–9.

Curriculum Vitae

Spillemansstraat 8
2140 Borgerhout (Antwerpen), België
+32 487 90 59 46
jorisvandenbossche@gmail.com
born in Borgerhout, 3 september 1988

Education

General education

2009-2011: Master of Science in Bioscience Engineering: specialization Environmental Technology, Ghent University

Master thesis: Modeling of BVOC emissions of a heterogeneous forest ecosystem

2006-2009: Bachelor of Science in Bioscience Engineering, University of Antwerp

2000-2006: Latin and Mathematics, H. Pius X - instituut Antwerpen

Employment

2012 - current: PhD in Bioscience Engineering, Ghent University

Title: *Towards high spatial resolution air quality mapping: a methodology to assess street-level exposure based on mobile monitoring*
The focus of the work is on the development of methodologies for mapping the urban air quality with mobile monitoring and low-cost sensors.
Research in collaboration with Vlaamse Instelling voor Technologisch Onderzoek (VITO) and KERMIT, Ghent University, under the supervision of Prof. dr. Bernard De Baets, ir. Jan Theunis and Prof. dr. ir. Dick Botteldooren.

Teaching-related activities

Scientific python courses and tutorials

ANTEA Scientific Python Workshop, March 9 and 11, 2015: two-day course for ANTEA given together with Stijn Van Hoey.

Introduction to Pandas, presentation given at PyData Paris, April 3, 2015, `https://github.com/jorisvandenbossche/2015-PyDataParis`.

Introduction to pandas, tutorial during EuroScipy, Cambridge, August 26-30, 2015 (link to schedule).

Tutorship

Het ontwerpen van een zwarte koolstof kaart van stadsregio Gent voor gebruik bij routing problemen. Annelies Van den Hove, master thesis (2015-2016).

Mobiele dataverzameling om de stedelijke luchtkwaliteit in kaart te brengen. Alexandra Van Wesemael, Fé Van der Elst, Inez Monteny & Lynn Van Ransbeke, bachelor project (2012-2013).

Scientific Output

Publications

Van den Bossche, J., Theunis, J., Elen, B., Peters, J., Botteldooren, D., & De Baets, B. (2016). Opportunistic mobile air pollution monitoring: a case study with city wardens in Antwerp. ATMOSPHERIC ENVIRONMENT, 141 408-421. `http://dx.doi.org/10.1016/j.atmosenv.2016.06.063`

Srbu, A., Becker, M., Caminiti, S., De Baets, B., Elen, B., Francis, L., ... & **Van den Bossche, J.** (2015). Participatory Patterns in an International Air Quality Monitoring Initiative. PloS ONE, 10(8), e0136763. `http://dx.doi.org/10.1371/journal.pone.0136763`

Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J. & De Baets, B. (2015) Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. ATMOSPHERIC ENVIRONMENT, 105 148-161. `http://dx.doi.org/10.1016/j.atmosenv.2015.01.017`.

Peters, J., **Van den Bossche, J.**, Reggente, M., Van Poppel, M., De Baets, B. & Theunis, J. (2014). Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. ATMOSPHERIC ENVIRONMENT, 92 31-43. `http://dx.doi.org/10.1016/j.atmosenv.2014.03.039`.

Campioli, M., Verbeeck, H., **Van den Bossche, J.**, Wu, J., Ibrom, A., D'Andrea, E., Matteucci, G., Samson, R., Steppe, K. & Granier, A. (2013). Can decision rules simulate carbon allocation for years with contrasting and extreme weather conditions? : a case study for three temperate beech forests. ECOLOGICAL MODELLING, 263 42-55. http://dx.doi.org/10.1016/j.ecolmodel.2013.04.012.

Presentations

KNMI City-Scale Air Quality Workshop, 19 October 2015, Oral presentation "Novel air quality monitoring with airQmap: DIY air quality mapping"

Peters, J., Srbu, A., **Van den Bossche, J.**, Ingarra, S., De Baets, B., Theunis, J., 2013. Development of a low-cost mobile sensor-system for participatory measurements of urban air quality. COST Action TD1105 EuNetAir 1st International Workshop, Open Satellite Workshop inside Transducers 2013, Barcelona, 20 June 2013. Barcelona.

Van den Bossche, J., 2013. Using Mobile Measurements for Air Quality Mapping: Methodological Issues. COST Action TD1105 EuNetAir WG3-WG4 Meeting, 6 March 2013, Duisburg.

Peters, J., **Van den Bossche, J.**, Theunis, J., & Van Poppel, M. (2013). The use of mobile air quality measurements to assess the spatial and temporal variability of urban UFP and BC concentrations. 4th EFCA Ultrafine Particles Symposium 2013.

Elen, B., Theunis, J., Ingarra, S., Molino, A., **Van den Bossche, J.**, Reggente, M., Loreto, V., 2012. The EveryAware SensorBox: a tool for community-based air quality monitoring, in: Workshop Sensing a Changing World. pp. 1–7.

Van den Bossche, J., Peters, J, Reggente, M., Botteldooren, D., Theunis, J., De Baets, B., 2012. The use of a mobile platform to assess the spatial and temporal variability of urban BC and UFP concentrations. Monitoring Ambient Air 2012, AAMG, London, 12-13 December 2012.

Posters

ISEE-EUROPE, Young Researchers Conference on Environmental Epidemiology, Barcelona, October 20-21, 2014. Poster presentations "Mapping exposure to urban air pollution based on mobile measurements by city guards in Antwerp, Belgium" and "Mapping spatial variation in exposure to urban air pollution using mobile measurements: development and validation of a methodology"

Monitoring Ambient Air 2012, AAMG, London, 12-13 December 2012. Poster presentation "City guards monitor urban air quality in Antwerp, Belgium".